Real time segmentation of heart sounds.

by

David Fourie

Thesis presented in partial fulfilment of the requirements for the degree of Master of Engineering (Electrical) in the Faculty of Engineering at Stellenbosch University

Supervisor: Dr. M.J. Booysen

December 2015

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: 2015/12/10 .....

Abstract

Real time segmentation of heart sounds.

D. Fourie

Department of Electrical and Electronic Engineering,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Thesis: MEng (Elec)

September 2015

The poor state of the healthcare system in South Africa has resulted in unacceptable high levels of infant mortality. Congenital heart disease is one of the main contributions to these high rates of mortality, with the cost of treatment and the availability of specialists being the driving factors. Computer aided auscultation is a technological solution to assist with the diagnosis of the disease. In its current form, computer aided auscultation is unsuitable for continuous patient monitoring.

The aim of this thesis is to develop an algorithm that will allow the existing methods of computer aided auscultation to work in real time so they can be used in patient monitoring. Existing methods of identifying the first and second heart sound are limited to offline processing. The algorithm developed in this thesis uses the correlation of the time-frequency coefficients of individual heart sounds to generate a feature vector for each heart sound that can be used to separate the sounds into different groups. To test the performance of the algorithm, 230 heart sounds from normal patients were first manually segmented and then processed with the algorithm. The noise sensitivity of the algorithm was also tested using generated heart sounds. Finally, the real time capability of the algorithm was tested.

The testing against sounds form normal patients resulted in a 84.2 % accuracy and an 84.4% hit rate. The synthetic testing showed the system starts to perform badly with a signal to noise ratio lower than -10db. The real time testing of the system showed that the algorithm is fast enough to be used in a real time environment. This thesis concludes that proposed algorithm is suitable for the detection of the first and second heart sounds in real time.

Uittreksel

Intydse segmentering van hartklanke

D. Fourie

Departement van Elektriese en Elektroniese Ingenieurswese,
Universiteit van Stellenbosch,
Privaatsak X1, Matieland 7602, Suid Afrika.

Tesis: MIng (Elek)

September 2015

Die toestand van die gesondheidstelsel in Suid-Afrika lei tot onaanvaarbare vlakke van kindersterftes. Oorerflike hartksiektes is een van die hoofoorsake van hierdie sterftesyfers, aangedryf deur die koste van behandeling en die tekort aan beskikbaarheid van spesialiste. Rekenaargesteundebeluistering is 'n tegnologiese oplossing wat help met die diagnose van hierdie kwaal. Huidiglik is rekenaargesteundebeluistering ongeskik vir aaneenlopende pasientemonitering.

Die doelwit van hierdie tesis in om 'n algoritme te onwikkel wat sal toelaat dat bestaande metodes van rekenaargestuendebeluistering intyds sal werk sodat gebruik kan word vir deurlopende monitering van pasiente. Bestaande metodes, wat aangewend word om die eerste- en tweede hartklanke te identifiseer, is beperk tot nie-intydse verwerking. Die algoritme wat in hierdie tesis ontwikkel is, gebruik die korrelasie van die tyd-frekwensie koeffisiente van individuele hartklanke om 'n eienskapsvektor vir elke hartklank te genereer, wat dan gebruik word om die hartklanke in verskillende groepe in te deel. Om die werkverrigting van die algoritme te toets, is 230 hartklanke van pasiente met normale harte eers per hand gesegmenteer en daarna met die algoritme verwerk. Die algoritme se bestandheid teen ruis is ook getoets deur gebruik te maak van sintetiese hartklanke. Uiteindelik is die intydse vermoë van die algortime getoets.

Die toetsing met normale hartklanke het 'n akkuraatheid van 84.2% en trefkoers van 84.4% opgelewer. Die ruisbestandheidstoetse het aangedui dat die stelsel sleg begin werkverrigting verloor met sein-tot-ruis-verhoudings van laer as -10dB. Die intydse toetsing het aangedui dat die algoritme vining genoeg is om gebruik te word in 'n intydse implementering. Die tesis maak die gevolgtrekking dat die voorgestelde algoritme geskik is vir die intydse identifisering van die eerste en tweede hartklanke.

Acknowledgements

I would like to express my sincere gratitude to the following people and organisations who have contributed to making this work possible;

- Dr. M.J. Booysen for his invaluable and patient support as study leader.
- Thys Cronje and Eugene Pretorius from Diacoustic medical devices for their financial support and useful guidance.
- My colleagues and the staff in the Electronic Systems Lab for accommodating a project that is completely different to the usual work of the lab.
- My family for their continuous encouragement.

Contents

List of Figures

List of Tables

Nomenclature

Medical abbreviations

CHD	Congenital Heart disease
VSD	Ventricular Septal Defect
ASD	Atrial Septal defect
PDA	Patent Ductus Arteriosus
CAA	Computer Aided Auscultation
ECG	Electrocardiogram

Signal processing abbreviations

STFT	Short Time Fourier Transform
CWT	Continuous Wavelet Transform
DWT	Discrete Wavelet Transform

Chapter 1

Introduction

Congenital heart disease (CHD) is the term for a range of conditions which are present in the heart at the time of birth. CHD has an incidence of 6/1000 live births for moderate to serious cases, and 60/1000 for all cases [3]. These rates are estimated to be relatively constant for difference countries [3]. The most common of these CHDs is the ventricular septal defect (VSD), patent ductus arteriosus (PDA) and the atrial septal defect (ASD) [3]. These conditions are usually identified by the murmur they produce during physical examination. This is problematic as heart murmurs occur in 80% of children, with the majority of them being innocent [4]. This leads to high rates of false positives when the patient is later examined with echocardiography [4]. The high rates of false positives increases the workload of specialists [5] and increases parental stress [6].

To address this problem of false positives, Computer Aided Auscultation (CAA) was developed to assist the primary care practitioner with the diagnosis. CAA uses signal processing and machine learning techniques to analyse the sound signals that emanate from the heart. Using CAA, medical practitioners are able to reduce the incidents of false positives [7].

The longterm continuous assessment of patients is commonly done to confirm the presence of a heart condition such as cardiac arrhythmia. These assessments are done using a portable device that records and analyses the Electrocardiogram (ECG) of the heart [8]. Whilst the ECG is capable of assessing certain conditions of the heart, it is unable to directly asses the movement of blood in the heart. This means that conditions such as PDA cannot be monitored automatically with an ECG. This gap forms the part of the basis of this thesis.

1.1 South African context

South Africa's healthcare system is not providing the level of care needed to the majority of its citizens [9]. 55% of South Africans rely on the public healthcare system for all their medical needs [10]. A system that is understaffed, underfunded and overburdened [9]. The system has to cope with the side effects of poverty such as malnutrition, low birth weight and HIV/Aids over and above the standard problems any healthcare system has to face such as non-communicable diseases [9]. The state of the healthcare system in South Africa is reflected in the mortality rates for children under 5 in Figure 1.1.

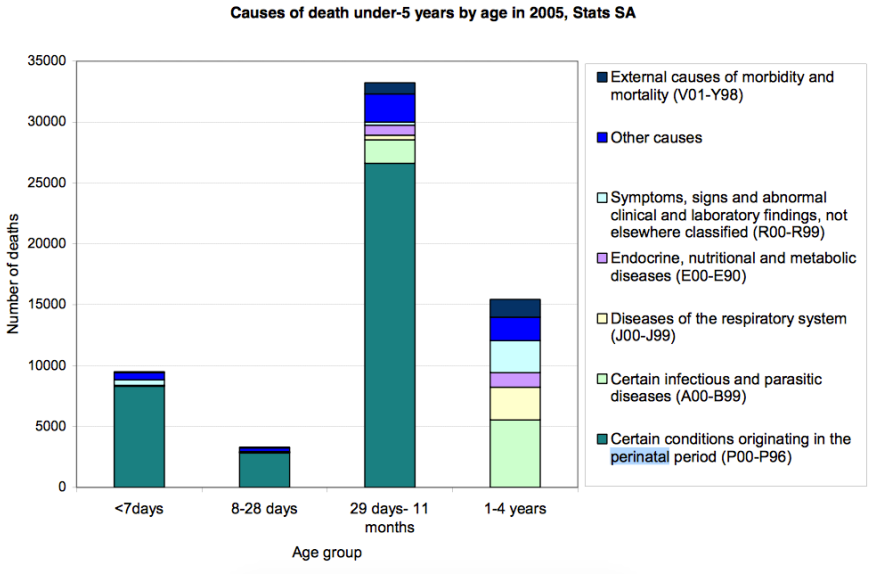

Figure 1.1: Figure showing the causes of mortality in children under 5 [1]

CHD in South Africa has an incidence of 0.6-0.8/1000 in liveborn children, which is about 11 000 per year [10]. CHD is a mostly treatable set of conditions with an 85% survival rate if there is adequate and timeous medical intervention [10]. In South Africa, the chances of getting the required intervention are low with only 25% of the children with CHD in South Africa receiving the treatment they need. The lack of treatment means that 3000 children die annually from CHD [10]. One of the reasons for this number is the number of referrals only represents a fraction of the total number of cases due to the condition being missed, misdiagnosed or identified too late [10].

1.2 Cardiac Auscultation

The method of auscultation was first practised during the Hippocratic period (460 to 370 BC). As the stethoscope was not invented yet, it was performed by a physician placing his/her ear to the patient's body. As well as being uncomfortable for the patient and physician, the sounds were difficult to hear. The development of the stethoscope by Laennec in 1816 enabled physicians to hear the sounds which the body makes more clearly, as well as making the process more comfortable for both patient and physician. As a result of the success of the stethoscope, auscultation is now part of the standard physical exam that is performed by medical practitioners [11].

Auscultation is performed during the physical examination of a patient to diagnose the function of a particular organ of the body. In cardiac auscultation, there are a number of sounds which are of interest to a medical practitioner. The different sounds are elicited from the body by moving the stethoscope to different parts of the chest and moving the patient to different positions [12].

Auscultation requires years of training by medical practitioners to become a useful tool for diagnosis. This combined with the increasing prevalence of echocardiography has resulted in a decline in the skill [13]. The effect on having medical practitioners with a lower effectiveness in auscultation will be a higher incidence of misdiagnosis and higher unnecessary cost for the healthcare system and patients.

1.3 Research gap

Computer aided auscultation has not been extended to work in real time for patient monitoring. In neonates with PDA for example there is a persistent murmur that indicates the presence of abnormal blood flow in the heart [14]. The monitoring of the condition is currently done using repeated echocardiograms, which require expensive equipment and employment of specialists. This drives up the cost of healthcare, both to the state and the patient. The murmur that is present in PDA is detectable with current offline auscultatory techniques [15]. The existing methods are easily adaptable to real time once the segmentation of heart sounds can be performed in real time. This is the basis for the research statements in this thesis.

1.4 Thesis statements and hypotheses

Thesis Statement 1:

Heart sounds can be segmented using short signals.

Hypothesis 1.1 The information in one heart sound is sufficient to tell it apart from another heart sound.

Hypothesis 1.2 A heart sound signal segmented with short signals is accurate.

Thesis Statement 2:

Heart sounds can be segmented in real time.

Hypothesis 2.1 The information in one heart sound can be analysed in a short amount of time.

1.5 Thesis overview

Chapter 2 presents an overview of the cardiovascular system, the types of sounds that medical practitioners look for during auscultation, a few of the conditions that are diagnosed with auscultation and an overview of the literature surrounding the process of computer aided auscultation.

Chapter 3 describes the signal processing and statistical methods used in this thesis.

Chapter 4 describes the methods used to prove the thesis statements as well as the testing procedure.

Chapter 5 documents how the methods were tested and their performance.

Chapter 6 concludes the work by validating the hypotheses in Chapter 1.

Chapter 2

Literature review

2.1 Cardiovascular system

The cardiovascular system is the system in the body responsible for delivering oxygen and removing carbon dioxide from the body. The organs and tissues which make up this system are the heart, the lungs and the blood vessels. The heart is responsible for the movement of blood though the system, the lungs are responsible for oxygenating the blood and removing carbon dioxide and the blood vessels are responsible for the transportation of blood. A diagram of this system is shown in Figure 2.1.

The focus of this thesis is the heart. The heart consists of two sides: the right side which is responsible for delivering de-oxygenated blood to the lungs and the left side which is responsible for delivering oxygenated blood to the rest of the body. Each side consists of two separate chambers called atria and ventricles. The atria are chambers in which blood collects from the the body and lungs. The ventricles are the chambers which contract and provide the pressure required to deliver blood to the rest of the body.

The atria and ventricles are connected by heart valves. The right atrium and ventricle are connected by the tricuspid valve and the left atrium and ventricle are connected by the mitral valve. The valves ensure the blood flows from the atria to the ventricles. The right ventricle is connected to the pulmonary artery via the pulmonary valve and the left ventricle is connected to the aorta via the aortic valve [2] [12] [16].

The positions of these valves are shown in Figure 2.2.

The path of blood through the heart explained below.

- The blood enters the heart through the vena cava and pulmonary vein. It fills the left and right atria.
- The atria contract and the blood enters the ventricles through the tricuspid and mitral valves.

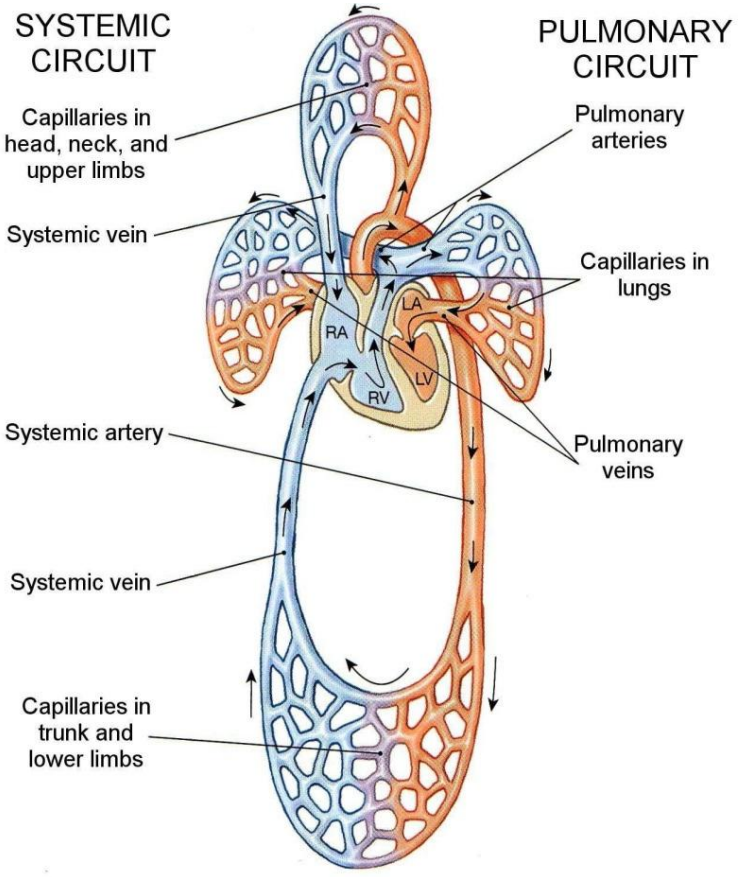

Figure 2.1: Diagram showing the movement of blood through the cardiovascular system [2]

- The ventricles contract and blood leaves the heart through the pulmonary and aortic valves into the pulmonary artery and aorta respectively.

2.2 Heart sounds

The two main sounds that emanate from the heart are caused by the rapid deceleration of blood that occurs when the valves in the heart close. These sounds are listed below.

First heart sound The first heart sound (S1) consists of the superposition of the sounds caused by the closing of the mitral and tricuspid valves. It marks the beginning of the systolic period and occurs when the ventricles contract [12].

Second heart sound The second heart sound (S2) occurs after the first heart sound. It consists of the superposition of the sounds caused by the closing

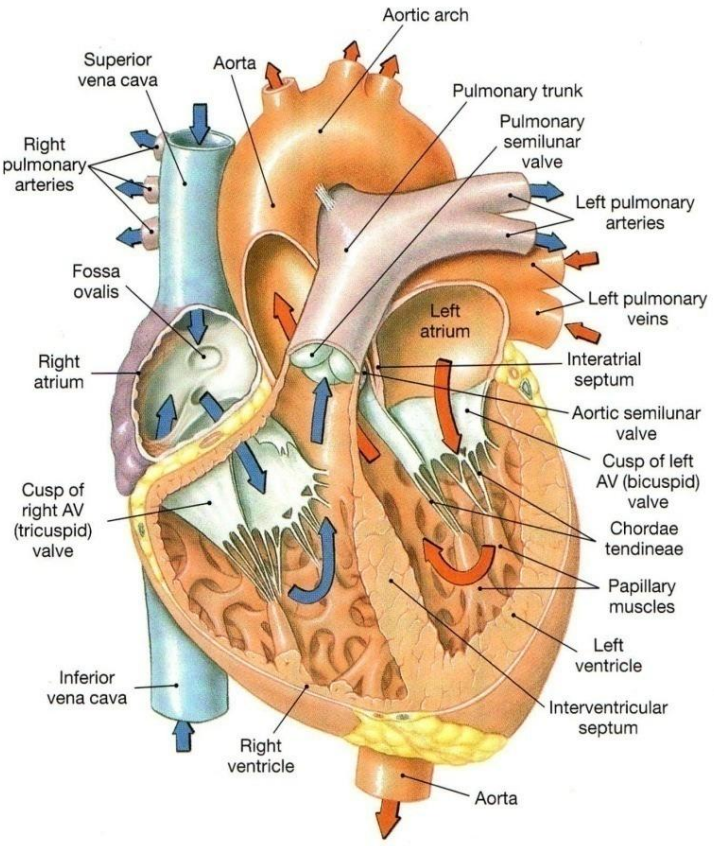

Figure 2.2: Diagram showing the anatomy of the heart [2]

of the near simultaneous aortic and pulmonary valves [16]. It marks the end of the systolic period and beginning of the diastolic period.

The other sounds that emanate from the heart have a variety of different causes.

Third heart sound The third heart sound (S3) is a low frequency sound that is sometimes heard after S2. It is caused by the oscillation of blood in the ventricles when they fill with blood. It is usually normal in patients under 40, but abnormal in patients over 40 years of age[17].

Fourth heart sound The forth heart sound (S4) is a low frequency sound that occurs just before S1. It occurs when blood enters into a still ventricle. It's presence can indicate the presence of certain heart conditions [18].

Murmur Murmurs in the heart occur when the otherwise lamina flow of blood in the heart undergoes turbulence. The turbulence creates vortexes, which are carried downstream. The area left by the vortexes is then rapidly filled with blood creating the sound. There are two classes of

murmurs, ejection murmurs and regurgitation murmurs. Ejection murmurs usually emanate from the aortic and pulmonary valves or the surrounding structures. Regurgitation murmurs occur when there is blood flow in the wrong direction through a valve. Some ejection murmurs are not indicative of a health condition and are innocent.

2.3 Cardiovascular disease detectable with auscultation

Some of the diseases that can be diagnosed with cardiac auscultation are listed below:

2.3.1 Atrial septal defect

An atrial septal defect is a congenital disease which causes blood to flow between the atria. When the symptoms are present, they are typical of cardiovascular disease (fatigue, exercise intolerance, shortness of breath). However most childhood cases remain asymptomatic until adulthood [19]. Initial diagnosis of an ASD is typically done using auscultation. The defect can cause a persistent split in the second heart sound as well as a soft systolic ejection murmur [19].

2.3.2 Ventricular septal defect

A ventricular septal defect is a congenital disease which causes blood to flow between the ventricles. The defect, if small, causes no symptoms and may close by itself. However if it is larger, it does cause symptoms which include fatigue, shortness of breath and exercise intolerance [20]. The diagnosis is typically performed during the physical examination, with a pansystolic murmur being the primary indicator of the condition [20].

2.3.3 Valvular stenosis

Stenosis of the heart valves is the narrowing of the valve opening. This narrowing decreases the blood flow through the heart as well as increasing the workload of the heart. These effects can cause the symptoms of exercise intolerance and shortness of breath [21]. They also lead to further complications that eventually cause heart failure. Aortic stenosis the most common form of valvular stenosis. The auscultatory signature of aortic stenosis is a systolic murmur [21].

2.3.4 Valvular regurgitation

Valvular regurgitation is flow of blood in the incorrect direction though a valve. This leads to an increase in the workload of the heart, which can eventually lead to heart failure [22]. The condition is initially asymptomatic, but eventually leads to heart failure if untreated. Mitral regurgitation is the most common form of regurgitation. The auscultatory signature of mitral regurgitation is a diastolic murmur [22].

2.4 Existing methods for cardiac auscultation

Most of the previous work in the field is related to the detection and identification of the first and second heart sounds in pre-recorded signals. This process is known as heart sound segmentation.There has been limited work done in the segmentation of heart sounds in real time.

2.4.1 Offline segmentation

2.4.1.1 Segmentation with signal envelope

The most basic method of heart sound segmentation uses envelope detection and physiological timing. The signal envelope is used to identify the location of heart sounds in the signal. Once the locations of the hearts sounds are found, they are identified by using the physiological timing differences between the first and second heart sounds.

The different types of enveloping methods were tested in [23]. The methods were successful in identifying the first and second heart sounds using their timing differences. The timing differences between the first and second heart sound decrease as the heart rate increases, thus this method is only suitable for patients with a low heart rate.

2.4.1.2 Segmentation with Neural Networks

A time delay neural network was used to identify the first heart sound using series of feature vectors generated from the energy of a continuous wavelet transform in [24]. The method used the electrocardiogram to identify S1 peaks in the heart sound signal or phonocardiogram (PCG). The time delay neural network was then trained to identify these peaks from the wavelet features. This method was successful in that it could reliably identify S1 peaks based solely on their time-frequency content. The main advantage of this method is that it can work on a short-time basis that could enable a system to perform the analysis while a caregiver is listening to the heart. The disadvantage of this method is that it relies on the availability of correctly segmented sounds to train the system.

2.4.1.3 Segmentation with wavelets

The physiological property of the heart that results in the second heart sound having more energy in the wavelet detail coefficient than the first heart sound, was used in [25] to identify lobes in the energy envelope of the approximation wavelet coefficients. A set of time domain criteria is used to remove lobes that do not correspond to the known time domain properties of heart sounds. The advantages of this method are it does; not require the use of machine learning methods and it can detect both the first and second heart sounds. The disadvantages of this method are its dependence on fixed frequency characteristics for segmentation and its dependence on time domain properties of heart sounds. This is problematic given the frequency content of heart sounds can vary between patients and pathologies and the reference. Using the time domain becomes problematic with tachycardia (abnormally fast heart rates) as the time differences between S1 and S2 start becoming smaller.

2.4.1.4 Segmentation with HMM

A Hidden Markov Model (HMM) was used to classify a sequence of mel-frequency cepstral coefficients (MFCC) of a specific sound [26]. To calculate the MFCC coefficients of a signal, the following procedure is followed: the power density spectrum is calculated for a specific segment of a signal; the spectrum is then projected onto the mel scale using an overlapping filterbank and the energy summed; the logarithm of the filterbank energies is calculated; finally the discrete cosine transform (DCT) of the energies is calculated, the coefficients of the DCT are the MFCC coefficients. The mel-scale is a logarithmic scale that is based on the perceived differences in human frequency identification. It attempts to scale the frequency spectrum in such a way that the differences between subsequent scales are constant. The advantage of this method is that it makes few assumptions about the time-frequency properties of the heart sounds. The disadvantage of this method is it requires prior learning, which makes its accuracy and generalisation a function of the quality and quantity of training samples. Since the HMM requires a sequence of values, it would be difficult to implement this method on a short time basis.

2.4.1.5 Segmentation with Empirical Mode decomposition

Empirical Mode Decomposition (EMD) and kurtosis statistics were used in [27] to determine the start and end of each heart sound. EMD is an iterative procedure that breaks a signal into a series of Intrinsic Mode Functions (IMFs). The IMFs are calculated to have two properties: the number of extrema and zero crossings must differ by at most one, and at any point, the mean of the envelopes defined by the local maxima and minima must be equal to 0. The set of IMFs is sorted to remove elements that do not correspond with known properties of heart sounds.The kurtosis descriptor is used to measure how much

the statistics of a signal segment match a Gaussian distribution. It is assumed that the statistics of a heart sound do not follow a Gaussian distribution, hence the kurtosis can be used to select heart sounds in the PCG. The advantage of this method is it makes few assumptions about the properties of heart sounds. The disadvantage is it would not work on a short time basis.

2.4.1.6 Segmentation with selectional regional correlation

The selectional regional correlations of time frequency data in [28], was used to classify different parts of the heart sound sequence. Selectional regional correlation is a pattern recognition method that uses the correlation of a template of known signal with an unknown signal to identify the known signal locations in the unknown signal. The advantages of this method are that it requires few assumptions about the signal it needs to classify and it can work on a short segment of data. The disadvantages of this method is it requires previous knowledge of the signals it needs to find.

2.4.2 Real time segmentation

Amplitude reconstruction was shown to be a feasible method for real time segmentation in [29]. The method uses amplitude thresholds to identify S1 and S2. Reported accuracy from the limited sample size was accurate, however the real time implementation of the method was only discussed.

2.4.3 Summary

The existing literature for heart sound segmentation systems does not sufficiently cover the analysis of heart sounds in real time. This makes them inefficient at analysing the variability of the heart sound signal. Furthermore, most of the existing literature depends on prior learning to train the system. The data for these systems comes from inconsistent biological sources and depends on the recording conditions such as the frequency response of the stethoscope and the recording location. This makes it difficult to create a generalised model of the system. The heart sound segmentation systems described in this chapter are summarised and compared in the Table 2.1.

Property / paper	[24]	[27]	[25]	[26]	[28]
S1 and S2	×	✓	✓	✓	✓
Few model assumptions	✓	✓	×	×	✓
Adaptive learning	×	✓	✓	×	×
Works in real time	×	×	×	×	×

Table 2.1: Table highlighting the contribution of this work against the state of the art.

Chapter 3

Development

This chapter describes the methods that are used to develop the real time segmentation system described in this thesis.

3.1 Enveloping methods

The envelope for a signal is a smooth, continuous function that outlines its extreme points. In heart sound segmentation, it is used to identify the position of heart sounds. The ideal signal envelope function for peak detection should do the following:

- Produce only one local maximum for each sound.
- Should minimise the width of envelope lobes to prevent overlap between sounds.
- Reduce the effect of noise

Three different methods of obtaining this function are detailed in the following section.

3.1.1 Hilbert Envelope

The Hilbert envelope is derived from the analytic representation of a signal [30]. The analytic representation of a real signal seeks to remove the negative frequency components. The result is a complex function which can be represented as the product of a real amplitude function and a phase function. This is shown below:

$$x_a(t) = x(t) + j\mathcal{H}(x(t)) \tag{3.1}$$

$$x_a(t) = \left|x_a(t)\right| \angle x_a(t) \tag{3.2}$$

$$x_a(t) = A(t)e^{j\Phi(t)} \tag{3.3}$$

where $x(t)$ is a real valued signal, $x_a(t)$ is the analytic representation of that signal, $A(t)$ is the magnitude of the analytic signal and $\Phi(t)$ is the phase of the analytic signal. An example of a signal and its Hilbert envelope is shown in Figure 3.1.

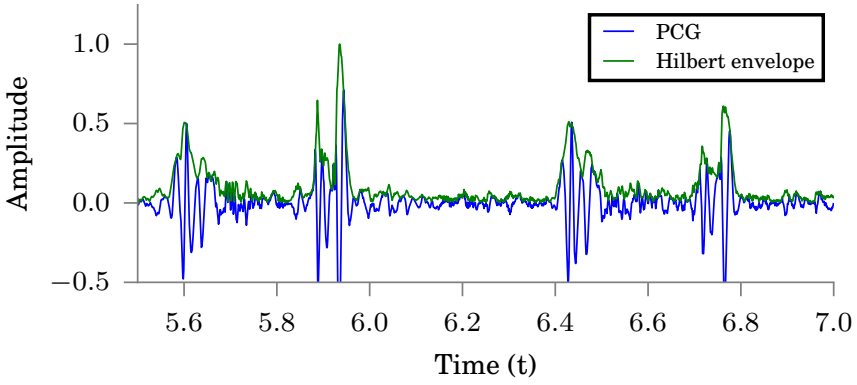

Figure 3.1: Graph showing the Hilbert envelope of a signal segment

The Hilbert envelope provides the precise definition of a signal envelope. However it does not necessarily produce a single local maximum for a given heart sound. It is therefore not ideal for finding a local maximum which would indicate the position of a heart sound.

3.1.2 Homomorphic Envelope

The homomorphic envelope described in [31] provides a smooth, continuous function that tracks the extrema of a signal. It is more impervious to noise from murmurs and external sources. The energy of the signal is modelled as the multiplication of two components. A high frequency component, which represents murmurs and noise, and a low frequency component which represents the components from S1 and S2. To separate the components, the logarithm of the signal amplitude is calculated. This process changes the multiplication of the components into addition, which can be filtered. The frequency components are assumed to remain separable by frequency content after the transformation. The transformed components are then filtered with a low pass filter to remove the higher frequency component. The resulting signal is then exponentiated to get the signal envelope. This process is shown below:

$$x(t) = a(t)f(t) \tag{3.4}$$

$$\log x(t) = \log a(t) + \log f(t) \tag{3.5}$$

$$L(\log x(t)) \approx \log a(t) \tag{3.6}$$

$$e^{L(\log x(t))} \approx a(t) \tag{3.7}$$

where $x(t)$ is the energy of the signal, $a(t)$ is the low frequency component of the energy, $f(t)$ is the high frequency component of the energy and L is the operator for a low pass filter. The cutoff frequency for the low pass filter L needs to be experimentally determined. Figures 3.2,3.3 and 3.4 show the effect of the cutoff frequency of L.

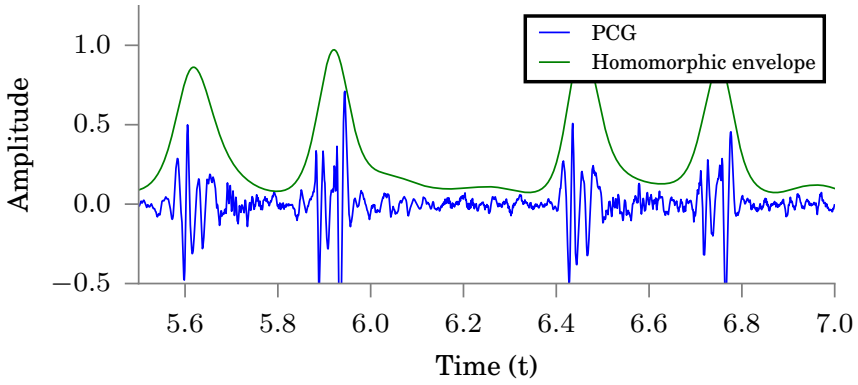

Figure 3.2: Graph showing the Homomorphic envelope of a signal segment using a lowpass filter of 5 Hz

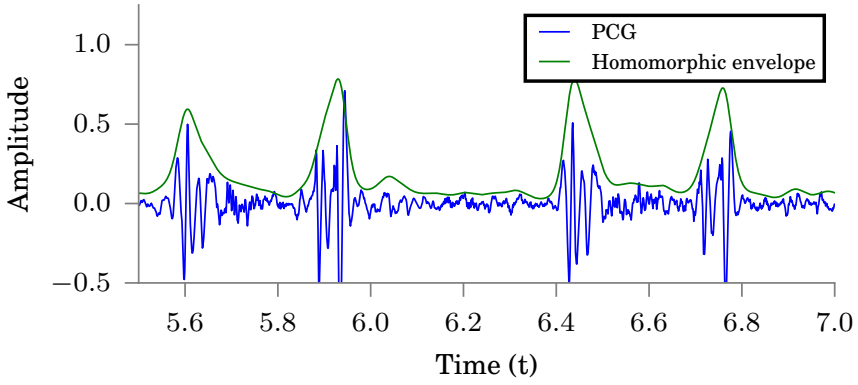

Figure 3.3: Graph showing the Homomorphic envelope of a signal segment using a lowpass filter of 10 Hz

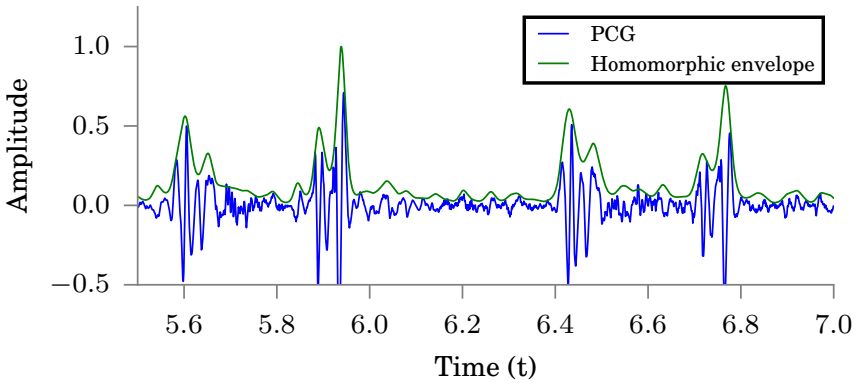

Figure 3.4: Graph showing the Homomorphic envelope of a signal segment using a lowpass filter of 20 Hz

The Homomorphic envelope produces a more desirable envelope than the Hilbert transform. Figure 3.3 only has one local maximum per sound and has relatively narrow lobes. However it is still sensitive to noise as a local maximum can be seen in some of the diastolic periods.

3.1.3 Shannon energy envelope

The Shannon energy envelope is derived from the Shannon entropy metric in information theory. The reason why the Shannon energy is used is that it performs as a soft amplitude threshold which emphasises the middle values of a normalised signal and de-emphasises loud and soft values. A comparison with other types of energy transformation is shown in Figure 3.5. The Shannon energy envelope is useful for a heart sound signal as the noise is typically soft (background noise) or loud (stethoscope movement). The calculation of the Shannon envelope is done by calculating the Shannon energy of the signal then averaging it over a fixed period, this is shown below:

$$x_s(t) = \frac{1}{N} \sum_{i=-N/2}^{N/2} x(t+i)^2 \times \log\left|x(t+i)\right| \tag{3.8}$$

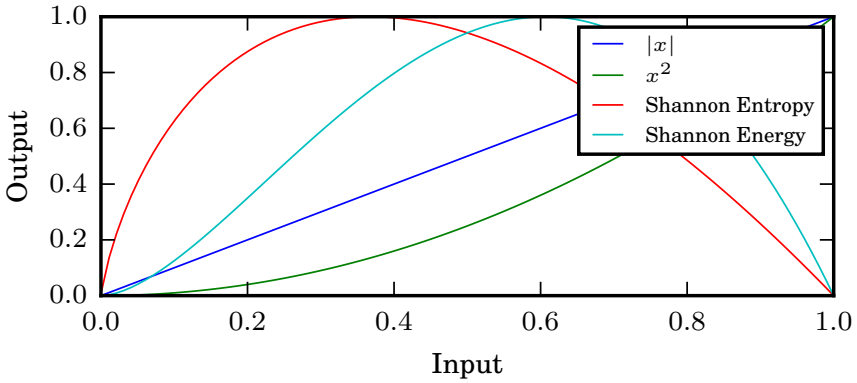

Figure 3.5: Graph showing the comparison of the different types of energy transforms.

The performance of the Shannon energy envelope is dependent on the value chosen for N. Using a signal with a sampling rate of 1000 Hz, the graphs in Figures 3.6, 3.7 and 3.8 were calculated to show the dependence of N.

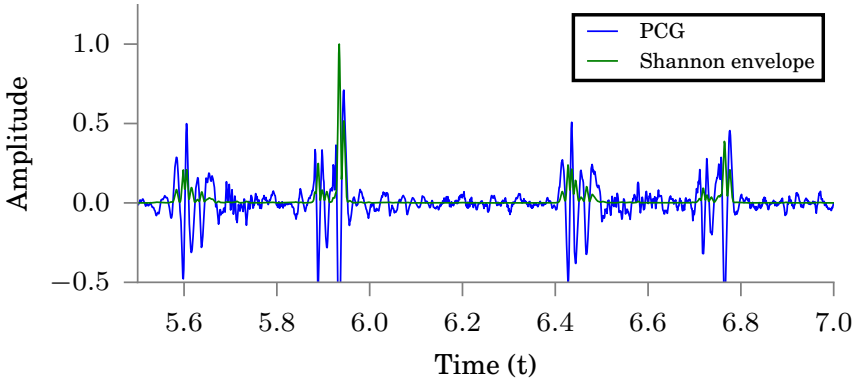

Figure 3.6: Graph showing the effect of $N = 5$ on the Shannon Envelope.

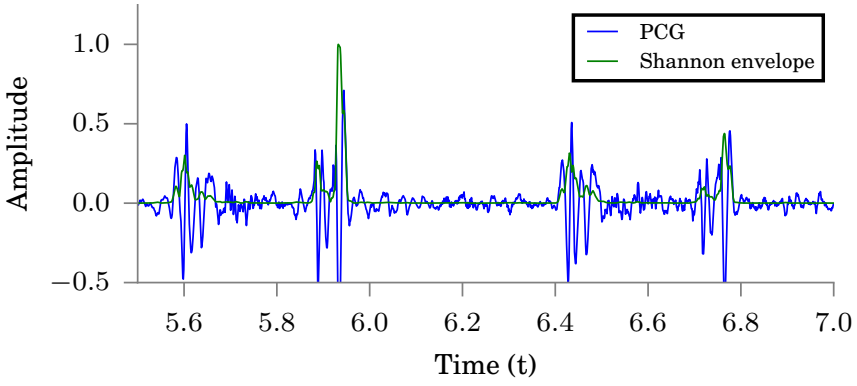

Figure 3.7: Graph showing the effect of $N = 10$ on the Shannon Envelope.

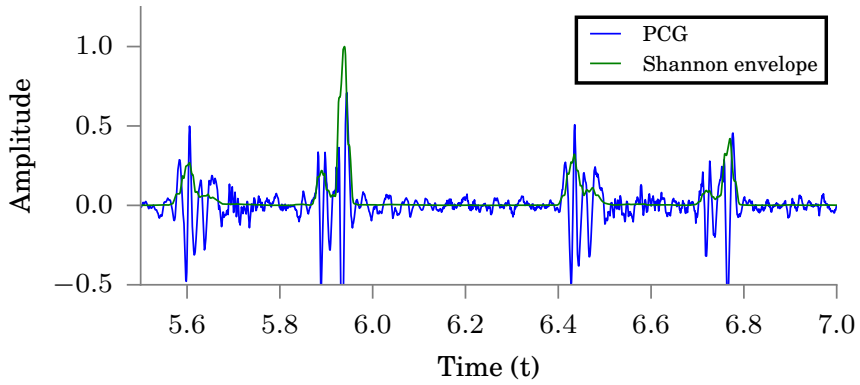

Figure 3.8: Graph showing the effect of $N = 20$ on the Shannon Envelope.

The envelope in Figure 3.8 provides narrow lobes, reduces the effect of the energy in the diastolic period and minimises the amount of local maxima per heart sound.

3.2 Peak detection

Peak detection is the process of finding all the local maxima in a function. In the analysis of heart sounds, the detection of peaks in the envelope of the signal is used to identify the location of heart sounds. From calculus, the local maxima and minima of a function can be found by finding the points where the derivative of the function is zero. This formulation works for continuous functions, however it does not necessarily work for discrete functions where the zero of the derivative could lie between the samples. The solution to this is to use the change of sign in the derivative. This is equivalent to calculating the second derivative. The direction of change is also used to identify the type of turning point. The process for identifying the local maxima of a signal is shown below:

$x_d[n] \leftarrow x[n+1] - x[n]$
$m \leftarrow 0$
for n=0 to N-1 **do**
  if SIGN($x_d[n]$)=1 and SIGN($x_d[n+1]$)=0 **then**
    $p[m] \leftarrow n$
    $m \leftarrow m + 1$
  end if
end for

where $x[n]$ is a discrete function with a length N, $x_d[n]$ is the discrete equivalent of the derivative of $x[n]$, p is a list containing the locations of the found local

maxima. The function SIGN is defined as follows:

$$\text{SIGN}(x) = \begin{cases} 1 \text{ if } x \geq 0 \\ 0 \text{ if } x < 0 \end{cases}$$

This method accurately identifies all the local maxima in the signal, however there are more local maxima in the signal than there are heart sounds, therefore the list of peaks needs to be filtered to select peaks that are of a large enough amplitude. The remaining peaks are then processed to resolve peaks that are within a specific distance of each other.

3.3 Time frequency methods

Time frequency methods are important to the analysis of heart sound signals. This is due to the non stationary nature of heart sounds. They are short sounds with changing frequency content. These conditions make frequency domain methods such as Fourier transforms ineffective for analysis. The time frequency methods that were examined in this thesis are detailed below:

3.3.1 Short Time Fourier transform

The Short Time Fourier Transform (STFT) uses the repeated application of Fourier transform to record the change in the frequency content over time. It is calculated by taking the Fourier transform of overlapping signal segments. This process is shown below:

$$\text{STFT}(\tau, f) = \int_{-\infty}^{\infty} x(t)w(t-\tau)e^{-2j\pi t}dt \tag{3.9}$$

Where τ is the time delay variable, f is the frequency variable, x is the time domain signal, w is a windowing function. The frequency accuracy of the STFT is defined by the width of the window function. A narrow window function will yield a transform that localises time accurately, but poorly localises frequency content, conversely, a wide window function yields poor time localisation, but good frequency localisation. This concept is explained by the Fourier Transform Uncertainty principle [32], which is stated below:

$$\left(\int_{-\infty}^{\infty}(t-t_0)^2\left|f(t)\right|^2 dt\right)\left(\int_{-\infty}^{\infty}(\omega-\omega_0)^2\left|\hat{f}(\omega)\right|^2 d\omega\right) \geq \frac{1}{16\pi^2} \tag{3.10}$$

where f is a normalised, square integrable function and $\hat{f}$ is the Fourier transform of the function. The principle states that the products of the second moments about zero of a normalised, square integrable function and its Fourier transform has to be greater than or equal to a fixed limit. The second moment

around zero of the norm of a function is a measure of the dispersal of energy in the function. The wider the dispersion, the larger the moment. The inequality only becomes an equality when f is a normalised Gaussian function. In the STFT, the dispersion of the function is controlled by the windowing function. Therefore the STFT only performs the analysis at a fixed resolution. This is a problem as the value of the frequency content changes with the frequency. For example, a low frequency sound usually doesn't need as much time localisation as a high frequency sound.

3.3.2 Wavelet analysis

Wavelet analysis is a generalisation of the Fourier transform which provides a more optimal time-frequency resolution than the STFT. There are two types of wavelet analysis, there is the Continuous Wavelet Transform(CWT) and there is the Discrete Wavelet Transform(DWT)

3.3.2.1 Continuous wavelet transform

The CWT is defined as the convolution of the daughter wavelet with the incoming signal. The daughter wavelet provides the optimal time-frequency function for a particular time and frequency.Using the variable names from [24], the wavelet transform of a function f at a time u and scale s is:

$$W_f(u,s) = \int_{-\infty}^{\infty} f(t)\Psi^*_{u,s}(t)dt \tag{3.11}$$

where $\Psi_{u,s}(t)$ is defined as the daughter wavelet of $\Psi(t)$ with the following relationship:

$$\Psi_{u,s}(t) = \frac{1}{\sqrt{s}}\Psi\left(\frac{t-u}{s}\right) \tag{3.12}$$

u and s represent the time delay and frequency scale variables.

The type of mother wavelet used for analysis is dependent on the application. For this thesis, the Morlet wavelet was chosen as it has a high time-frequency resolution [24]. The Morlet wavelet is defined as a time-shifted gaussian function modulated with a complex exponential. This function is shown below:

$$\Psi(t) = \pi^{-\frac{1}{4}}\left(e^{-j\omega_0 t} - e^{-\omega_0^2/2}\right)e^{-t^2/2} \tag{3.13}$$

The value of ω_0 is used to change the time-frequency resolution trade off. The affect of ω_0 is shown in Figures 3.9 and 3.10. The higher frequency scaling functions in 3.10 will have better frequency resolution, but are wider in time so they will have poorer time resolution. The scaling functions in Figure 3.9 have a lower frequency resolution, but have a higher time resolution as they are narrower. The effect the value of omega on the analysis of heart sounds is

shown in Figures 3.11 and 3.12. Figure 3.11 has a higher time resolution so is able to resolve time differences between signal components with more accuracy, but does not have sufficient resolution to resolve frequency component data. Figure 3.12 has a higher frequency resolution than 3.11 so more frequency information is visible at a cost of time resolution.

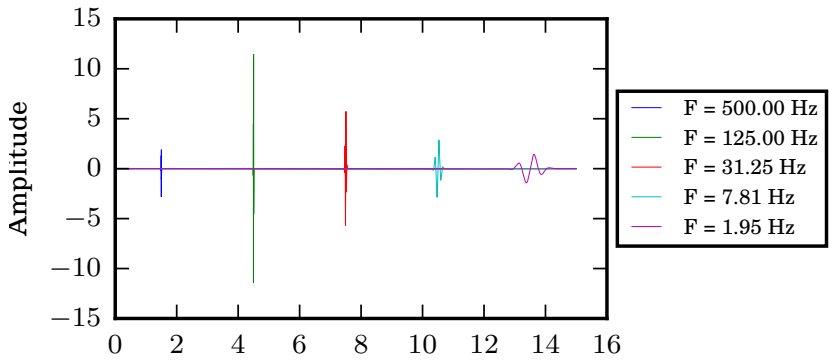

Figure 3.9: Graph showing the effect of $\omega_0 = \pi$ on scaling functions for different frequencies

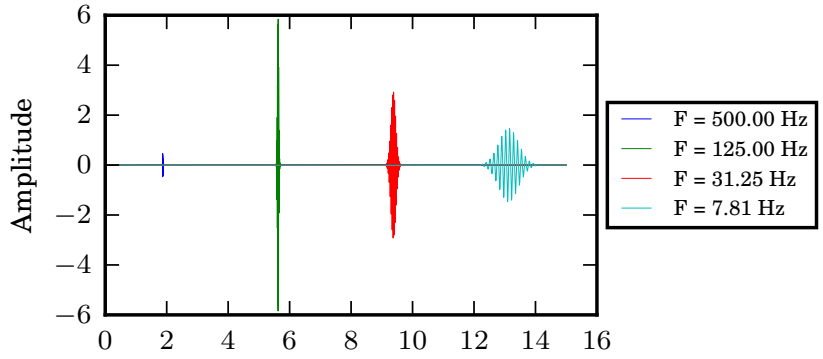

Figure 3.10: Graph showing the effect of $\omega_0 = 5\pi$ on scaling functions for different frequencies

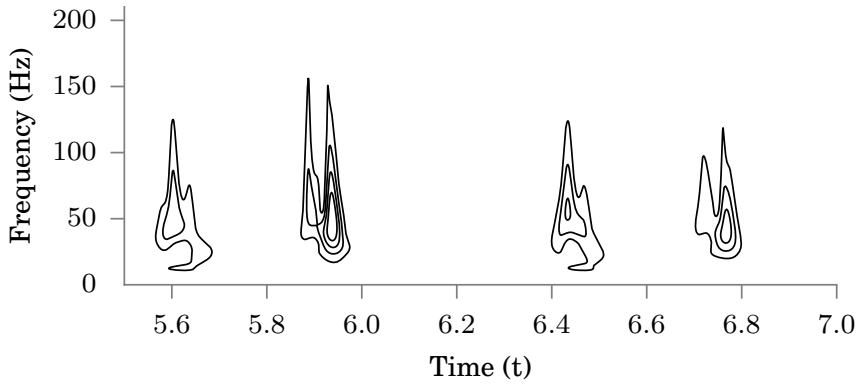

Figure 3.11: Contour plot showing the effect of $\omega_0 = \pi$ on the CWT of heart sounds

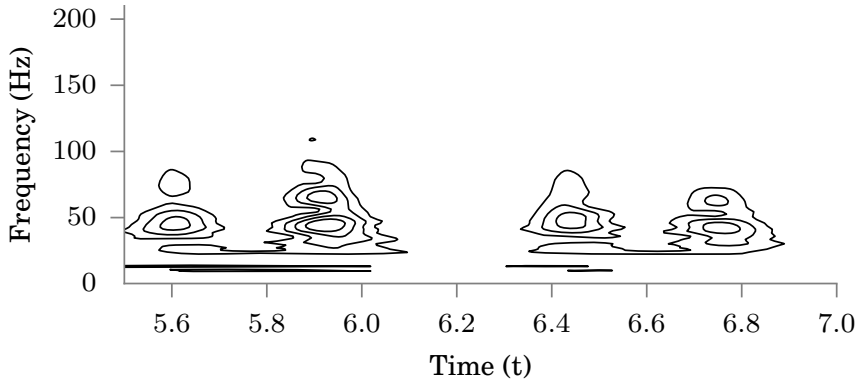

Figure 3.12: Contour plot showing the effect of $\omega_0 = 5\pi$ on the CWT of heart sounds

3.3.2.2 Discrete wavelet transform

The Discrete Wavelet Transform is an adaption of the CWT to use discretely sampled wavelets. The transform is implemented using a set of discrete filters arranged in a filter bank with downsampling by two between the levels of the bank. The outputs of the filter banks are the coefficients of the wavelets. The set of filters consists of a bandpass $h[n]$ and a low pass filter $g[n]$. The coefficients from the bandpass filter are called the detail coefficients and the coefficients from the low pass filter are called the approximation coefficients. This is shown below:

$$d[n] = \sum_{m=-\infty}^{\infty} x[m]h[2n - m] \tag{3.14}$$

$$a[n] = \sum_{m=-\infty}^{\infty} x[m]g[2n - m] \tag{3.15}$$

where $x[n]$ is the input signal, $d[n]$ are the detail coefficients, $a[n]$ are the approximation coefficients.

This is shown graphically in Figure 3.3.2.2.

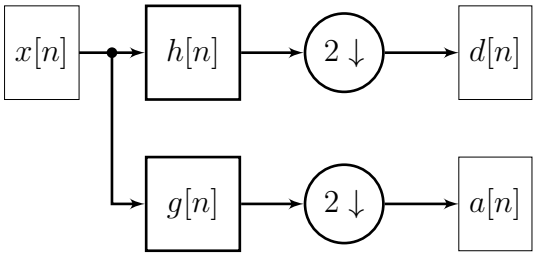

Figure 3.13: Block diagram of one level in a DWT filter bank.

These levels are cascaded to analyse different levels of frequency content with each downsampling process halving the level of frequency analysis. The low pass filter ensures that the downsampling process does not remove any frequencies. The cascaded filterbank with three levels is shown in Figure 3.3.2.2.

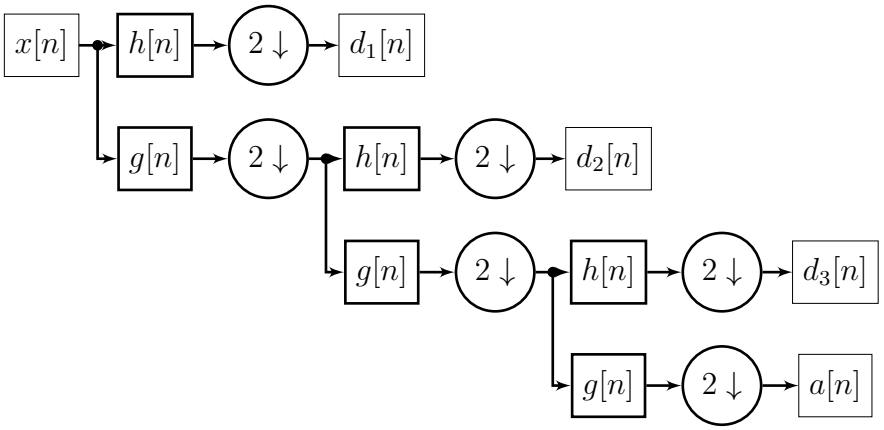

Figure 3.14: Block diagram of three levels of a DWT filterbank.

The time-frequency resolution is adjusted by changing the type of wavelet. The Daubechies family of wavelets are commonly used in time/frequency analysis. Figures 3.15 and 3.16 show the differences between the different types of wavelets on the DWT.

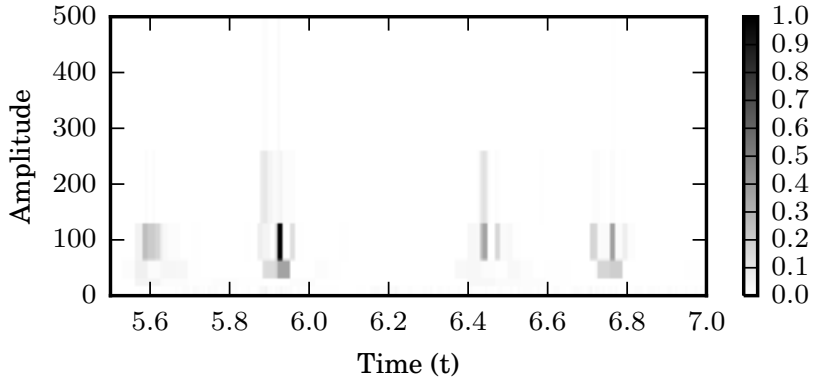

Figure 3.15: 2D heatmap of the distribution of DWT coefficients for the *db*6 wavelet.

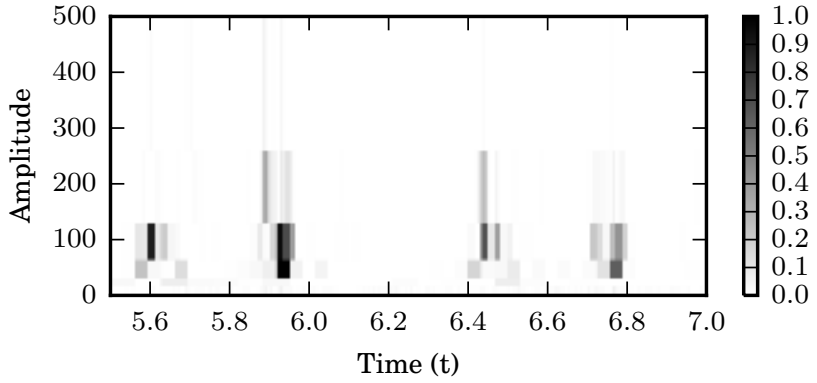

Figure 3.16: 2D heatmap of the distribution of DWT coefficients for the *db*3 wavelet.

Both of the figures show the change in size of the time-frequency analysis windows. Figure 3.16 has a higher time resolution than figure 3.15, this allows for the shorter, higher frequency components to become visible.

3.3.3 Instantaneous frequency

The frequency component of a signal can be analysed using the analytic representation of the signal. In analytic signals, the instantaneous frequency of a signal is represented by the derivative of the phase. The normal way of calculating the phase of a signal is to use the arctan function. This becomes problematic with digital signals as the phase becomes periodic and is bounded

between $-\pi$ and π. To get around this problem, the following derivation used.

$$\omega[n] = \dot{\rho}[n] \tag{3.16}$$
$$x_a[n] = A[n]e^{j\rho[n]} \tag{3.17}$$
$$\rho[n] = \mathrm{Im}[\ln(x_a[n])] \tag{3.18}$$
$$\omega[n] = \mathrm{Im}[\frac{d}{dn}(\ln(x_a[n]))] \tag{3.19}$$
$$\omega[n] = \mathrm{Im}[\frac{\dot{x_a}[n]}{x_a[n]}] \tag{3.20}$$
$$\omega[n] \approx \mathrm{Im}[\frac{x_a[n+1] - x_a[n]}{x_a[n]}] \tag{3.21}$$
$$\tag{3.22}$$

where $\omega[n]$ is the instantaneous frequency and $\rho[n]$ is the phase of $x[n]$. A frequency sweep is shown in Figure 3.17 to demonstrate the ability of the Hilbert instantaneous frequency to track changes in the signal frequency. The boundary effects at the beginning and end are caused by the lack of a windowing function in this signal.

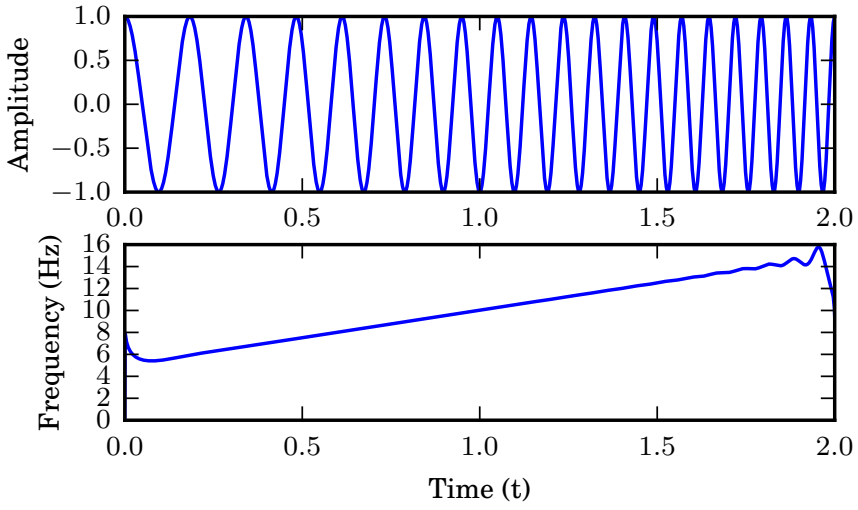

Figure 3.17: Figure showing a frequency sweep from 5 Hz to 15 Hz

The instantaneous frequency of a signal segment from a heart sound is shown in Figure 3.18. The high frequency noise in the signal makes this type of analysis difficult.

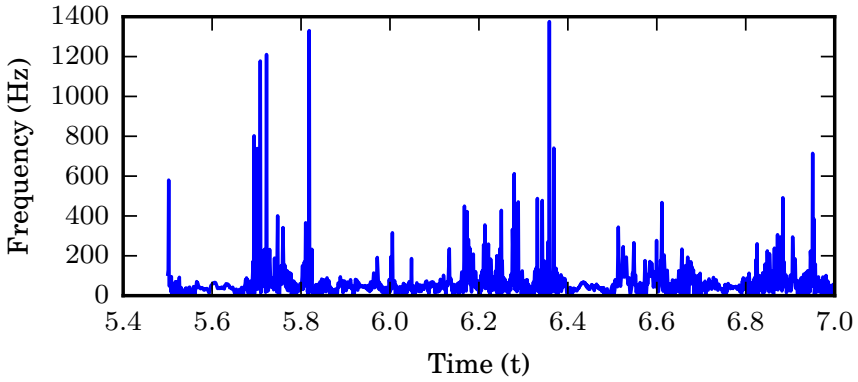

Figure 3.18: Plot showing the instantaneous frequency content of heart sounds

3.4 Statistical clustering

K-means clustering is a statistical method of grouping together a set of vectors into a specified k number of clusters [33]. Each cluster is associated with a mean that determines how the date points are assigned to each cluster. The algorithm iterates over the following steps until the assignment of vectors does not change between iterations.

Assignment Each vector is assigned a cluster based on minimising the Euclidean distance from each particular vector to the centroid of the cluster.

Update New centroids calculate by finding the mean of all the vectors assigned to a particular cluster,

3.5 Summary

This section introduced the various signal processing and statistical techniques of analysis used in this thesis.

Chapter 4

Method

This chapter introduces the segmentation method and testing procedure used in this thesis.

4.1 Inspiration for method

The coefficients from a continuous wavelet transform contain sufficient information to identify S1 from S2 [24]. The current technology utilises machine learning to separate the heart sounds with these coefficients. This method does not take advantage of the inherent repetition of the different heart sounds. To take advantage of this property, correlation is used, specifically the correlation of the continuous wavelet transform coefficients. The CWT was used over the DWT as the system was implemented on a desktop computer which could calculate the higher resolution CWT of a heart sound sufficiently quickly for real time processing.

4.2 Algorithm

There are two algorithms that this thesis introduces. An algorithm that segments pre-recorded sounds and an algorithm that segments sounds that are not pre-recorded.

4.2.1 Offline sound segment identification

The following assumptions are made about the sounds that are emanate from the heart:

- The time-frequency content for S1 and S2 cover similar regions in the domain
- The time frequency content for S1 is different than that of S2

- The selectional regional correlation of one sound with another yields a maximum when the two sounds are similar.

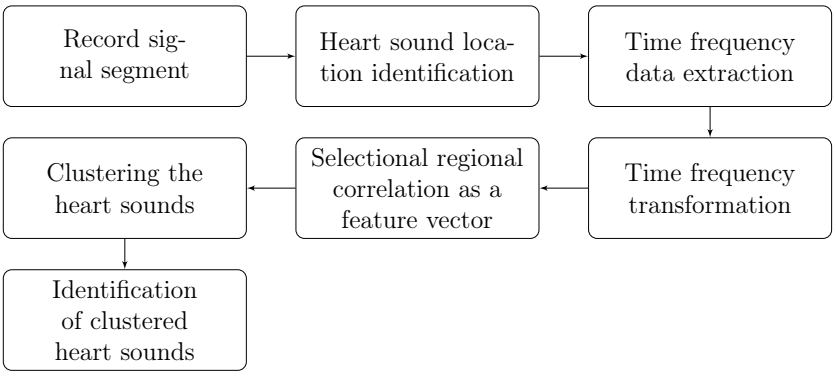

Figure 4.1: Flow diagram showing the layout of the S1 - S2 detection system

4.2.1.1 Heart sound location identification

To identify the origins of all the sounds in a signal, the location of each sound is required. This is achieved by identifying the peaks in the Shannon envelope of the signal. The Shannon envelope of a heart sound segment is shown in Figure 4.2.

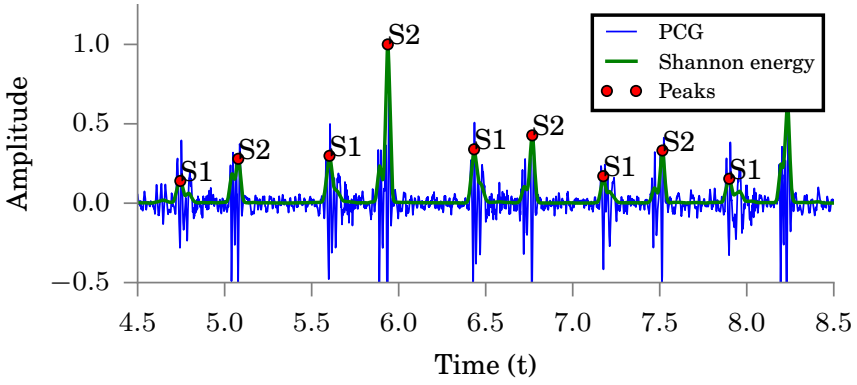

Figure 4.2: Graph showing the Envelope, PCG and identified peaks.

4.2.1.2 Time frequency data extraction

The CWT for each sound is calculated from a segment of signal around each located peak. The segment length is calculated from the width of the peaks in the Shannon envelope. The magnitude of the CWT of the signal in Figure 4.2 is shown in Figure 4.3.

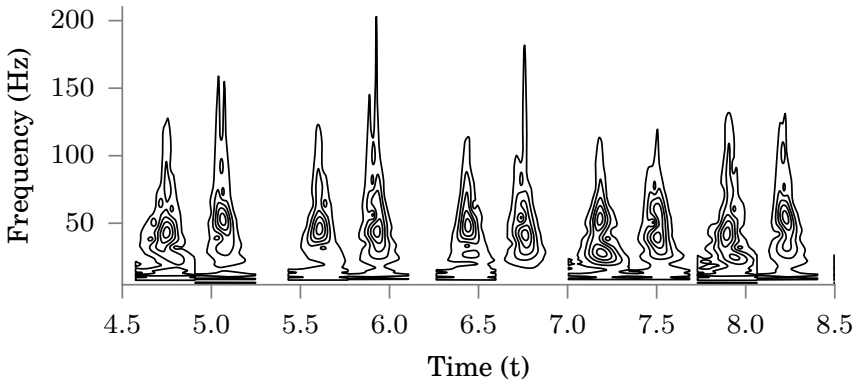

Figure 4.3: Graph showing wavelet coefficients of the signal.

4.2.1.3 Time frequency transformation

Correlation of the time frequency components is used to group together similar sounds. To make this process more accurate the following properties are needed of the signals to be correlated:

- The amplitude of the signals needs to be constant, otherwise higher valued signals bias the correlation
- The effect of noise needs to be minimised

To achieve these properties, the Shannon energy envelope is calculated for the coefficients. The Shannon energy is used for similar reasons as it was in the envelope. It reduces the effect of high and low amplitude noise. The transformed coefficients are then thresholded to values of 1 or 0 using high and low thresholds. This is to eliminate the effect of amplitude on the correlation as well as emphasising a specific shape and position for the sound in the domain. The result of the transformation and correlation is shown in Figure 4.4.

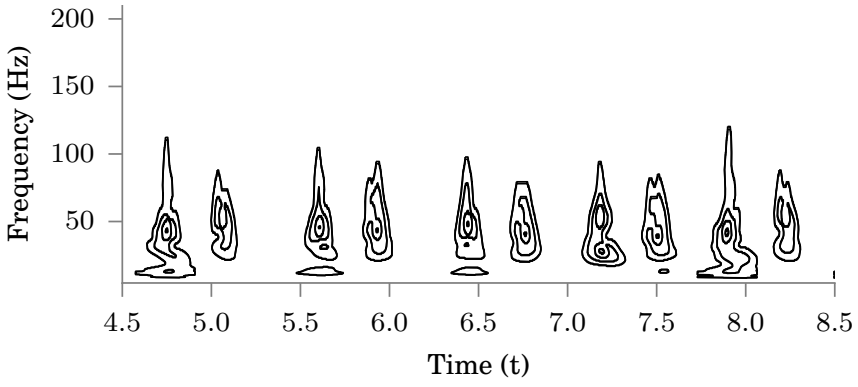

Figure 4.4: Graph showing transformed wavelet coefficients of the signal.

4.2.1.4 Selectional regional correlation as a feature vector

To segment the located sounds, the selectional regional correlation for each sound is calculated with respect to all the other sounds in the signal. The time frequency correlation for two sounds is calculated by multiplying the components of the two sounds together and summing the result. The correlation for a particular sound is expected to increase when the compared sound is similar and decrease when the sound is dissimilar. The patterns of increasing and decreasing levels of similarity are shown in Figure 4.5.

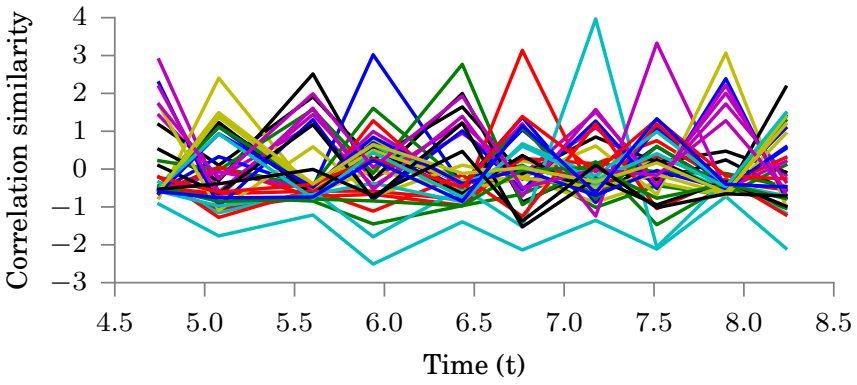

Figure 4.5: Graph showing the correlation of the transformed wavelet coefficients.

4.2.1.5 Clustering the heart sounds

The selectional regional correlation for a particular sound is used as a feature vector for a machine learning system. Each dimension in the vector represents how similar the sound is with another sound. The heart sounds are then clustered using the k-means clustering algorithm. The clustered correlations are shown in Figure 4.6.

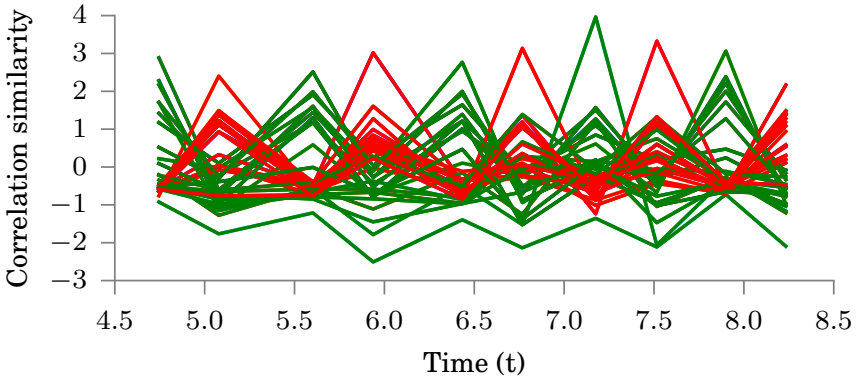

Figure 4.6: Graph showing the grouped correlations of the signal.

4.2.1.6 Identification of clustered heart sounds

The clustered correlations are then used to separate the heart sounds into two groups. These groups will each contain a number of sounds that are similar to each other. The next step is to identify the source of each group of similar sounds. The time difference between successive S1 and S2 sounds is smaller than time difference between successive S2 and S1 sounds. Using this information, the groups are identified as either S1 or S2 based on the average transaction time between the identified groups. Whilst this method is similar to the method in 2.4.1.1, the comparison is made between groups of similarly sounding heart sounds as opposed to ungrouped sounds.

4.2.2 Online segmentation

The algorithm in section 4.2.1 is adapted to real time usage by using a template of identified sounds, which is generated during a short setup stage of a few seconds, to group an identified sound instead of using k-means. An overview of the methods is shown in Figure 4.7.

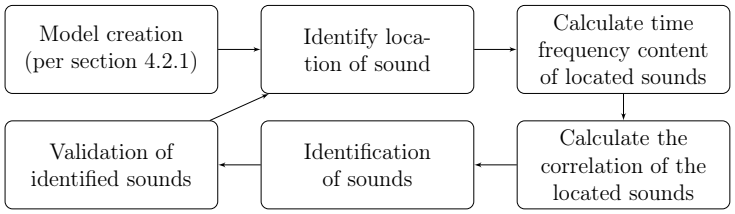

Figure 4.7: Flow diagram showing the layout of the real time system

4.2.2.1 Model creation

To identify a heart sound in real time, the system first needs to generate templates of S1 and S2 sounds with which an unknown sound can be correlated. The templates are constructed from the transformed time frequency coefficients of identified S1 and S2 an initial recording (5 seconds minimum). The templates are arranged in an alternating pattern of groups, namely S1 and S2. With the templates in a pattern, the correlation of the templates is calculated and the mean value of the correlation in each group is found. This is the equivalent of performing the k-means algorithm on the correlations.

4.2.2.2 Identify location of sound

The real time analysis of the signal is performed on overlapping 2 second chunks of data with a refresh rate of 1 second. These intervals guarantee that at least one sound is captured in the chunk as the average heart rate is 60 to 100 beats per minute. In each 2 second chunk of data, the Shannon envelope

is calculated and the peaks are found. The peaks represent the location of a potential heart sound.

4.2.2.3 Calculate time frequency content of located sounds

The time frequency content of each identified sound is calculated and transformed in the same way as subsections 4.2.1.2 through 4.2.1.3.

4.2.2.4 Calculate correlation of identified sounds

The identified sounds are correlated with the S1-S2 template to generate feature vectors to identify the sounds.

4.2.2.5 Identification of sounds

The sounds are identified by calculating the Euclidean distance from their feature vectors to each of the means calculated in section 4.2.2.1. The group with the smallest distance is chosen as the label for the sound being identified.

4.2.2.6 Validation of identified sounds

To increase the accuracy of the system, a simple method of validation is used. The method consists of only using pairs of sounds that correspond to the physiological process. Therefore the only sound pairs that are used are the S1 to S2 pair and the S2 to S1 pair. Any other pair is considered to be an error in identification.

4.3 Synthetic testbed

The testing of any algorithm is essential to determine its effectiveness, specifically testing how an algorithm copes with a low signal to noise ratio. This type of testing is difficult for a heart sound segmentation algorithm as the signals are inherently noisy and do not have established models for the different sounds. There are also problems concerning the type of noise in the signal. There is firstly uncorrelated white noise which would represent background noise from the body and recording environment. This type of noise is relatively easy to test with, as it involves adding generated white noise to a signal that is considered noise free then adjusting the amplitude of the noise to obtain the desired signal to noise ratio. The other type of noise that is associated with heart sounds is the correlated noise that arises from murmurs and S3 and S4 heart sounds. This type of noise is more problematic for a segmentation algorithm as well as being more difficult to test. This section shows the development of a synthetic testing platform using a sum of fitted Gaussian modulated sinusoidal functions.

4.3.1 Signal synthesis

The synthetic heart sounds are generated from a previously segmented heart sound signal. For each detected heart sound in the signal, a sum of gaussian sinusoids is fitted to the sound using the following algorithm:

- Set initial parameters to a low frequency
- Use least squares fitting to change the parameters of the function to minimise the error between the original function and the fitted function
- Subtract the fitted function from the original function, increase the initial frequency and repeat.

The Gaussian modulated function and its parameters are shown below:

$$G(t, A, f_0, \sigma, t_0, \phi) = A\cos(2\pi f_0(t - t_0) - \phi)e^{-(\frac{t-t_0}{\sigma})^2} \tag{4.1}$$

The process is shown in Figures 4.8,4.9,4.10 and 4.11.

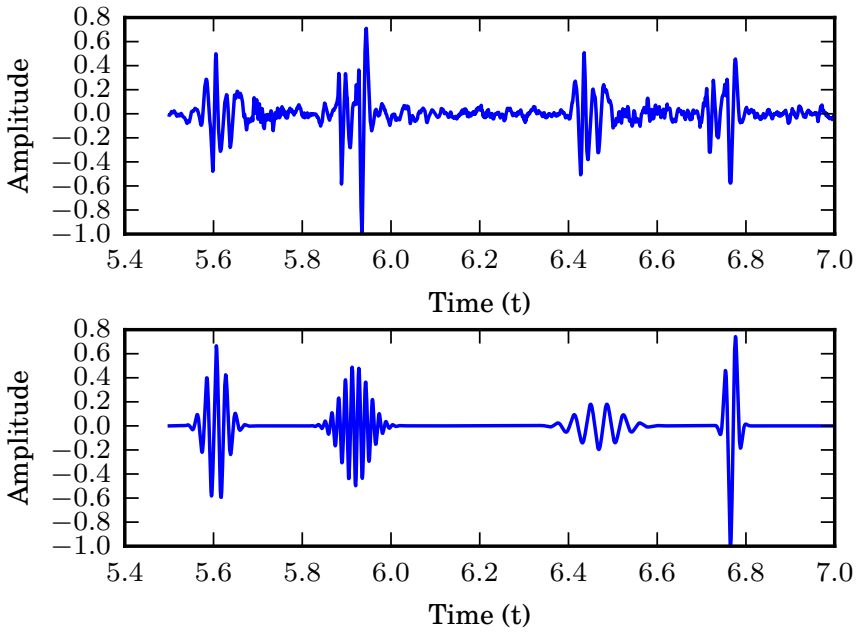

Figure 4.8: Graphs showing the original signal and the first approximation.

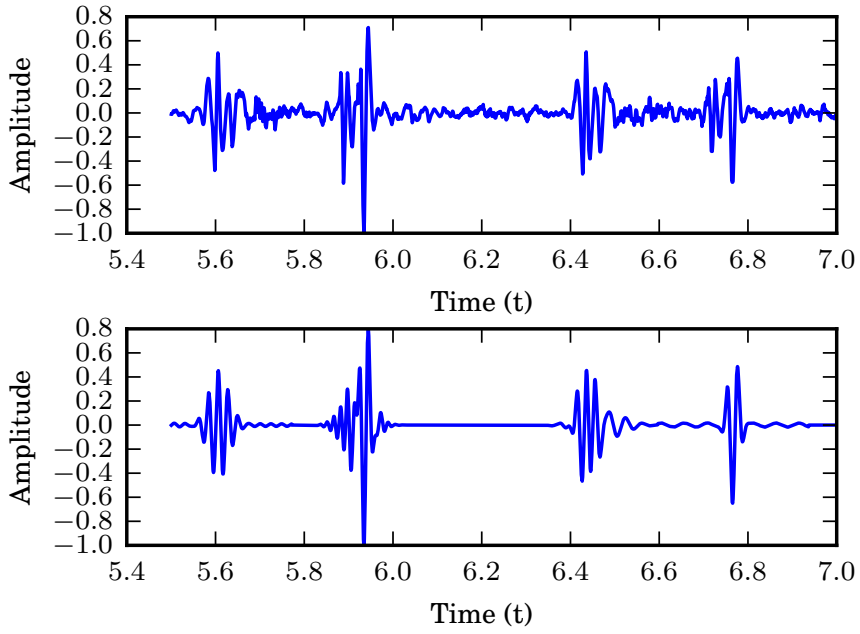

Figure 4.9: Graphs showing the original signal and the second approximation.

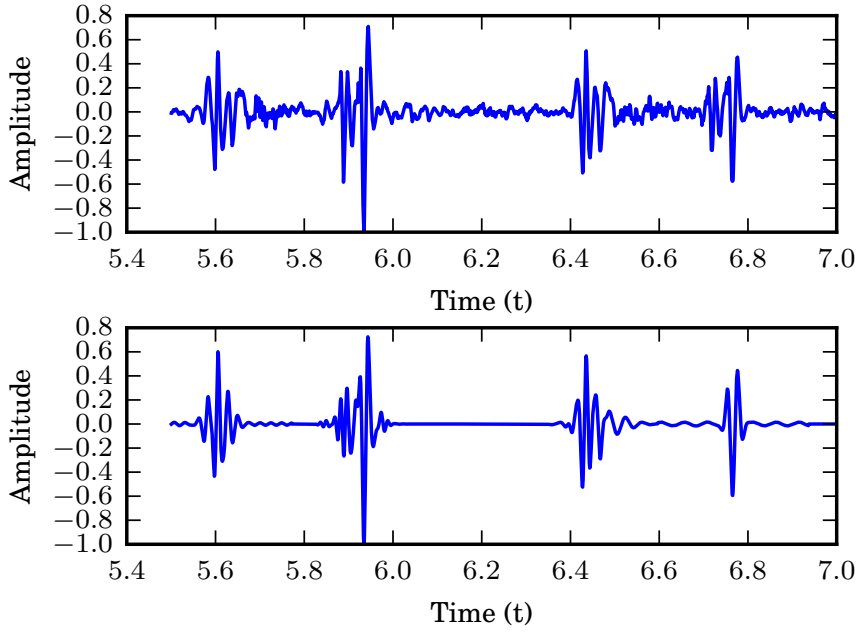

Figure 4.10: Graphs showing the original signal and the third approximation.

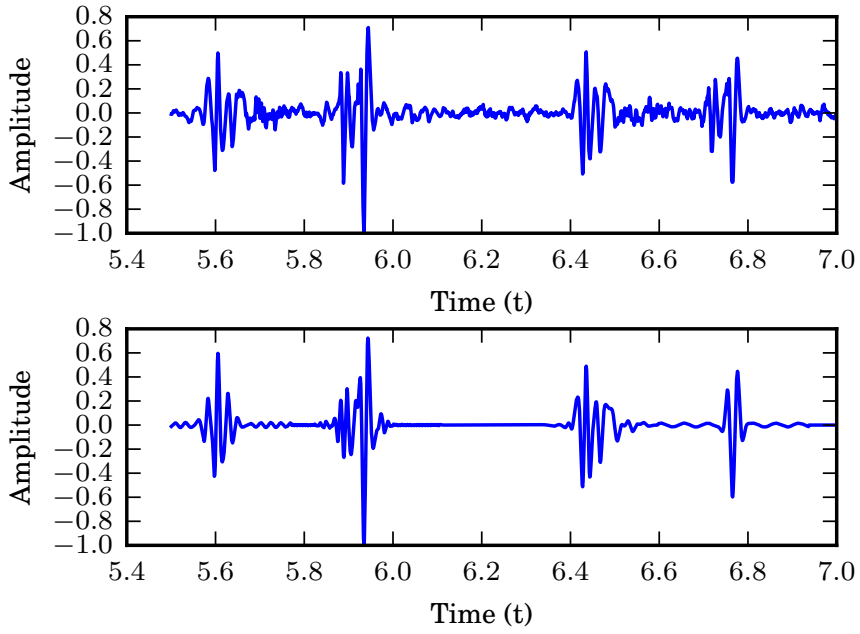

Figure 4.11: Graphs showing the original signal and the final approximation.

The parameters for each level of approximation in each heart sound are used to build a random distribution from which new parameters can be drawn from. This is done to approximate the random nature of biological signals.

4.3.2 Signal generation

The test signals are generated by drawing the parameters for the Gaussian functions from the random distributions. The process for generating the signals starts with determining the positions of the heart sounds based on the heart rate. Once the positions are calculated a heart sound is generated for that position by using the values from the random distribution. To ensure that heart sounds in their respective groups have a degree of similarity, the distributions for the two approximations with the highest energy are modified to have a lower standard deviation. These two approximations are then treated as the part of the signal which contains the information required for segmentation. The rest of the approximations are treated as additive noise. Gaussian white noise is then added to the signal to simulate background noise from the body. Three different generated signals for different heart rates are shown in Figures 4.12, 4.13 and 4.14.

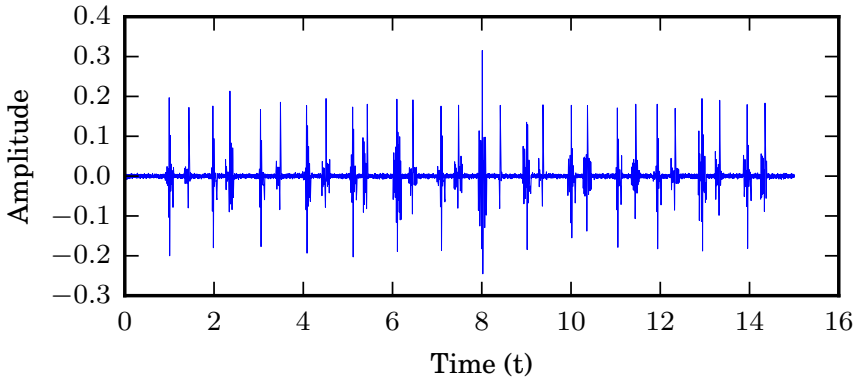

Figure 4.12: Graph showing a 60 BPM generated heart signal.

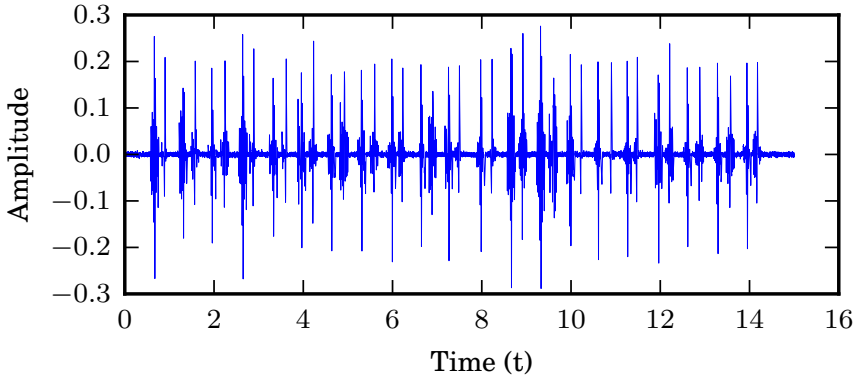

Figure 4.13: Graph showing a 90 BPM generated heart signal.

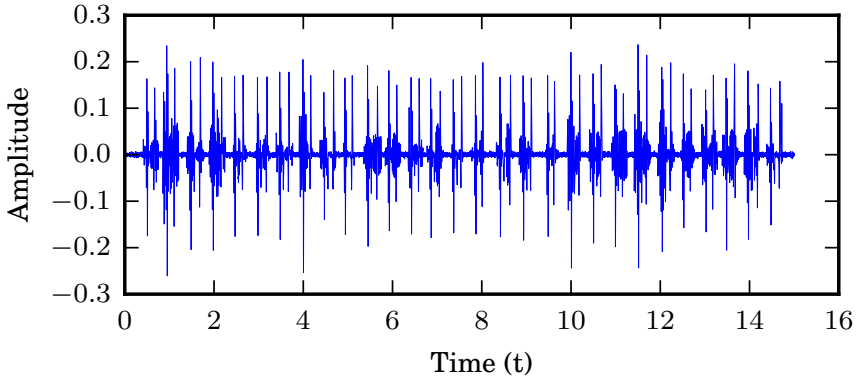

Figure 4.14: Graph showing a 120 BPM generated heart signal.

4.4 Summary

This chapter introduced a heart sound segmentation method that can detect the first and second heart sounds in real time. It also introduced a method for generating heart sound signals with a controllable signal to noise ratio.

Chapter 5

Testing

This chapter documents the testing done on the algorithms developed in chapter 4. The three types of testing done in this chapter are offline, synthetic and online testing.

5.1 Offline testing

This section documents the testing of the system using prerecorded heart sounds from test subjects. The sounds were recorded at the Red Cross children's hospital under ethical conditions. The dataset was captured from the four areas of auscultation in 230 healthy subjects.The signals were recorded at 22 kHz and were each 15 seconds long. The prerecorded signals were manually segmented with the assistance of an ECG and the results saved in a database. Each of the signals were segmented using the offline method described in section 4.2.1.

5.1.1 Offline testing results

The results of segmentation are shown in Table 5.1:

	Manually identified S1	Manually identified S2	Manually identified other
System identified S1	2444	508	15
System identified S2	334	2168	18

Table 5.1: Table to show results of the system implemented in this thesis.

5.1.2 Offline testing evaluation

The two measures used to evaluate the performance of the algorithm are the heart sound hit rate and the heart sound accuracy. The heart sound hit rate refers to the amount of sounds correctly identified out of the total number of sounds of that type in the signal. The hit rate gives a measure of how sensitive the system is to detecting a heart sound in a signal. For these results, the hit rate is:

$$S1_{HR} = \frac{2444}{2444 + 334} = 87.9\% \tag{5.1}$$

$$S2_{HR} = \frac{2168}{2168 + 508} = 81.0\% \tag{5.2}$$

The heart sound accuracy rate refers to the amount of sounds that were correct out of the sounds that were identified. This gives a measure of how much the system's results can be trusted. For these results, the accuracy rate is:

$$S1_{AR} = \frac{2444}{2444 + 508 + 15} = 82.3\% \tag{5.3}$$

$$S2_{AR} = \frac{2168}{2168 + 508 + 18} = 86.1\% \tag{5.4}$$

$$\tag{5.5}$$

The average implementation time on a dual core i5 laptop is broken down below:

CWT calculation : 0.09 seconds

CWT correlation : 0.038 seconds

Feature vector calculation : 0.0128 seconds

k-means : 0.0003 seconds

Total : 1.36 seconds

5.1.3 Offline testing discussion

The evaluation of the algorithm shows the system is capable of accurately detecting the majority of heart sounds in the database with an average hit rate of 84.4% and an average accuracy rate of 84.2%. These performance rates prove that the algorithm in this thesis is capable of segmenting heart sounds reasonably accurately. When testing the performance of segmentation algorithms, it is difficult to construct a fair test that accurately simulates real world performance. The difficulty comes from the source of the testing data as there is no control over the conditions that generate the data, other than the

selection of the data itself. The dataset used in this testing for example could contain 15 % of fundamentally unusable sounds, if this was the case, having a hit rate of near 85 % would actually mean the algorithm is working very well. The problem is it is impossible to objectively obtain the unusable sounds. One way to reduce the effect of this is to apply some selection criteria to the signal, such as a visual threshold. This would increase the apparent performance of the algorithm as there would be less undetectable sound. This process might lead to incorrect conclusions about the performance of the algorithm as the data are cherry picked. On the other hand, no algorithm is capable of producing results from data that for all intents and purposes is junk, so including it would unfairly decrease the apparent performance as well. The only real to get around this problem is to use a large set of data to test the algorithm on and hope that there is an equal distribution of "perfect" signals and junk signals. This method has its problems as well as to manually verify such a large dataset is unfeasible. This is what motivated the use of synthetic testing.

The implementation time break down indicates that the calculation time for the feature vector is fast enough to be implemented in real time as the distance between subsequent heart sounds is significantly more than 12 milliseconds.

5.2 Synthetic testing

The testing in Section 5.1 only shows how the system performs against real signals with little control on the quality. To test the segmentation ability of the algorithm in controlled, noisy environments, the generated signals with a known signal to noise ratio from Section 4.3.1 are used. To set the signal to noise ratio, a synthetic signal is generated with separate noise and heart signal components. The SNR for this signal is calculated and used to calculate a constant multiple for the heart signal which will set the SNR to the desired level. This calculation is shown below:

$$SNR_{old} = 10 \times \log\left(\frac{\sum_{i=0}^{N} x_s[i]^2}{\sum_{i=0}^{N} x_n[i]^2}\right) \tag{5.6}$$

$$c = 10^{\left(\frac{SNR_{des} - SNR_{old}}{20}\right)} \tag{5.7}$$

$$x_s[i] = c \times x_s[i] \tag{5.8}$$

where SNR_{old} is the signal to noise ratio of the generated signal, x_s is the heart signal component, x_n is the noise signal component and SNR_{des} is the desired signal to noise ratio. Two types of noise are tested in this section, correlated noise and uncorrelated noise. Correlated noise refers to the additional Gaussian modulated components that are drawn from a random distribution. Uncorrelated noise refers to Gaussian white noise. The testing in this section tests the algorithm against both types individually and in combination.

The testing procedure test the algorithm against a range of signals with different SNRs. At each chosen level of SNR, 50 different signals are generated, segmented and evaluated, with the metrics from Subsectionsect:eval used to evaluate the performance. Twenty different levels equally spaced were chosen between -200dB and 5 dB, yielding a total of 2000 different sounds tested for each heart rate. Three different heart rates were tested.

5.2.1 Synthetic testing interpetation

The graphs in Subsections 5.2.2, 5.2.3 and 5.2.4 show that noise starts hindering the accuracy at -10dB and the hit rate at -10dB, regardless of the type of noise added. The graphs also indicate that the sensitivity to noise increases with the heart rate, also regardless of the type of noise added. The graphs show the type of noise that has the most effect on the segmentation ability is uncorrelated noise, whilst correlated noise has the least effect. The addition of correlated noise to uncorrelated noise did not present a significant change to the segmentation ability of the system.

The synthetic testing provides insight into where this algorithm will be useful. The robustness against the correlated noise means that the algorithm is robust against the time-frequency variation that occurs between heart sounds. The sensitivity to uncorrelated noise does not affect the effectiveness of the algorithm as this type of noise is easy for a health care provider to identify and correct.

5.2.2 Uncorrelated noise testing results

Four examples of a generated signal with this type of noise at different levels are provided in Figures 5.1, 5.2 and5.3

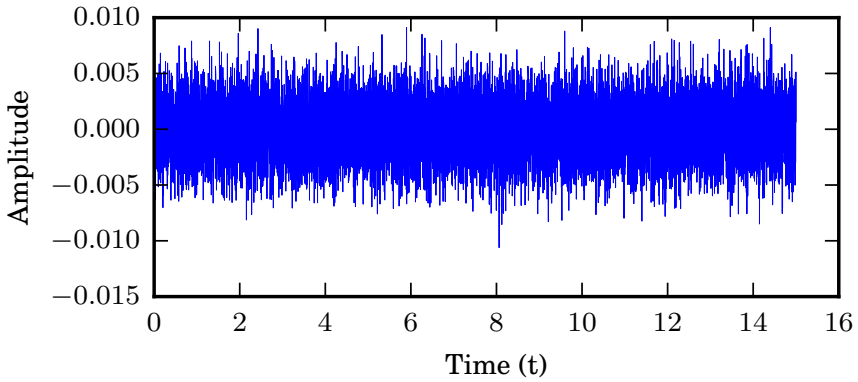

Figure 5.1: Graph showing a generated signal with uncorrelated noise with a SNR of -50dB.

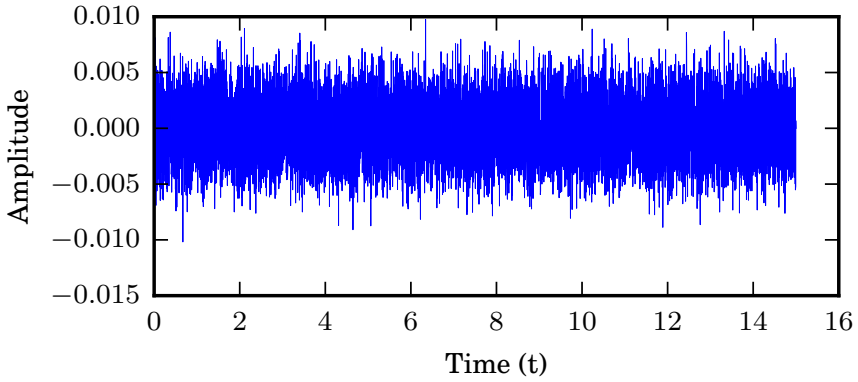

Figure 5.2: Graph showing a generated signal with uncorrelated noise with a SNR of -15dB.

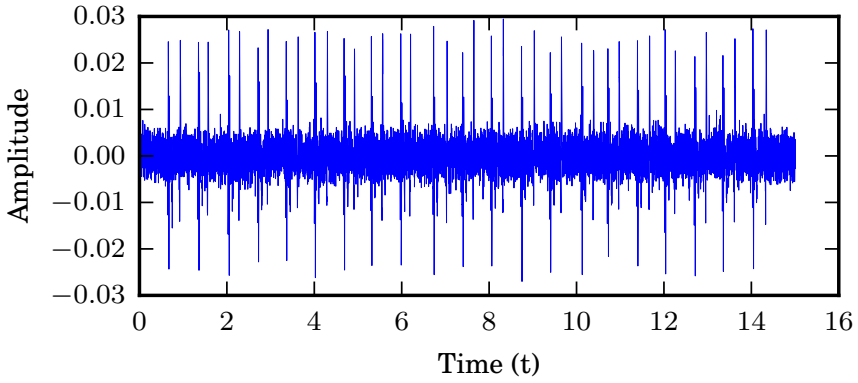

Figure 5.3: Graph showing a generated signal with uncorrelated noise with a SNR of 5dB.

The results for the testing at three different heart rates are shown in Figures 5.4, 5.5, and 5.6.

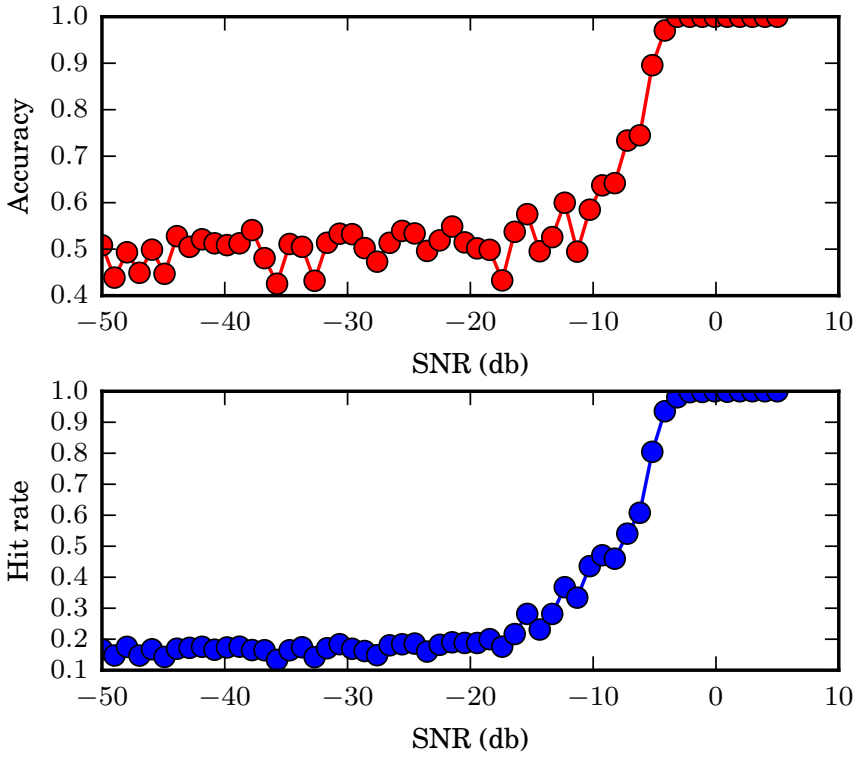

Figure 5.4: Graph showing the relationship between SNR and performance at 60 BPM.

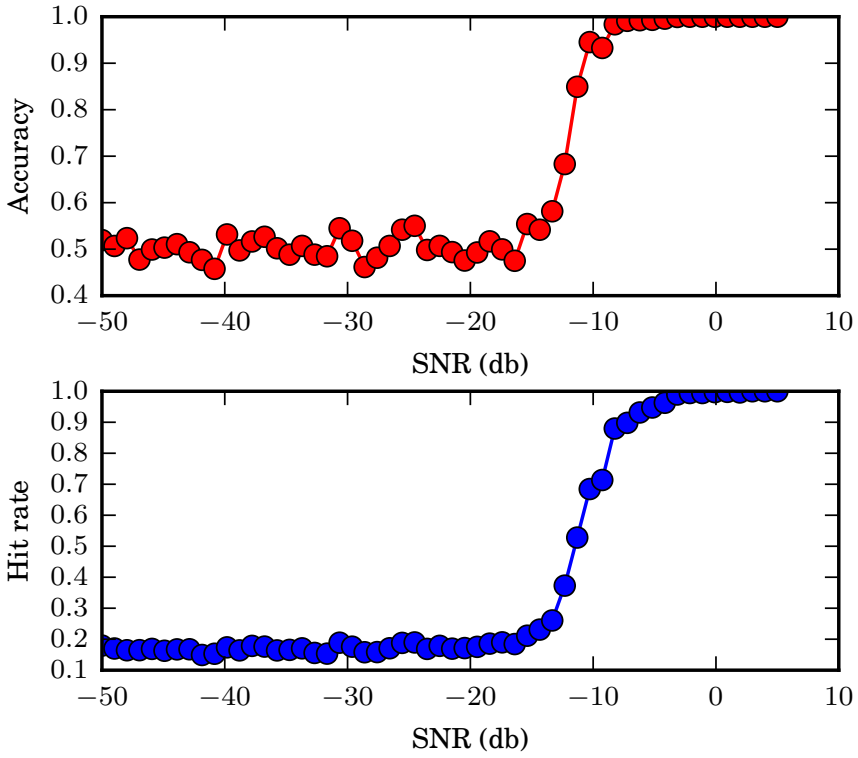

Figure 5.5: Graph showing the relationship between SNR and performance at 90 BPM

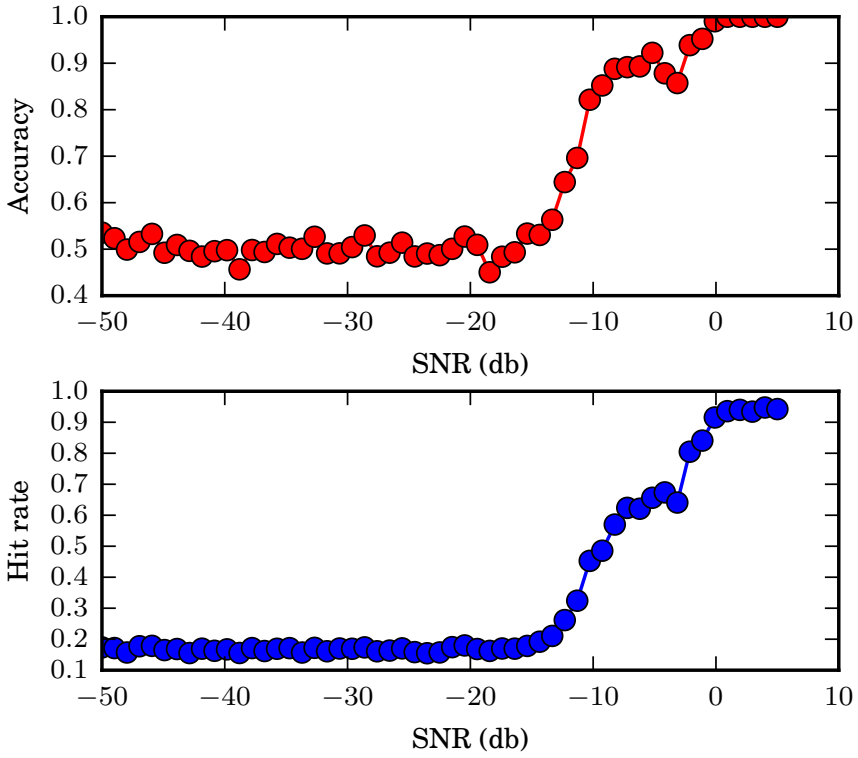

Figure 5.6: Graph showing the relationship between SNR and performance at 120 BPM.

5.2.3 Correlated noise testing results

Four examples of signals with different levels of correlated noise are shown in Figures, 5.7, 5.8 and 5.9.

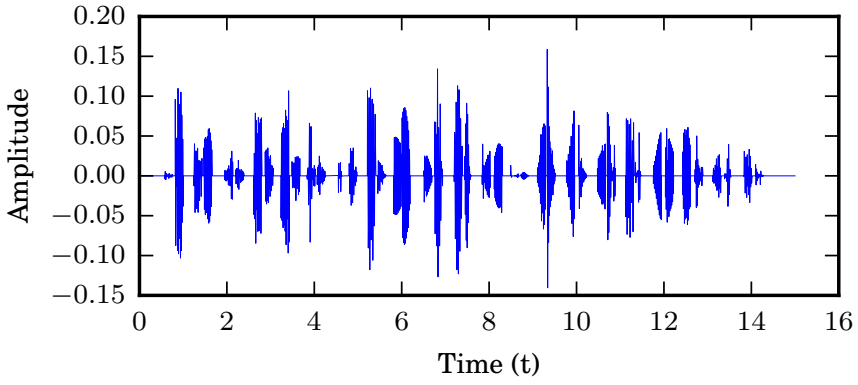

Figure 5.7: Graph showing a generated signal with correlated noise with a SNR of -50dB.

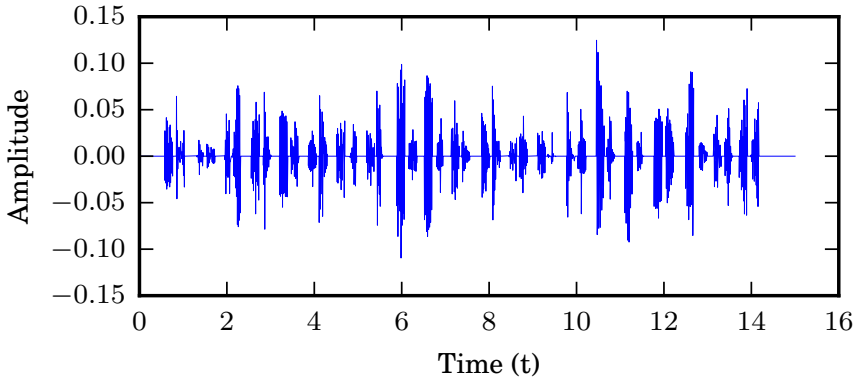

Figure 5.8: Graph showing a generated signal with correlated noise with a SNR of -15dB.

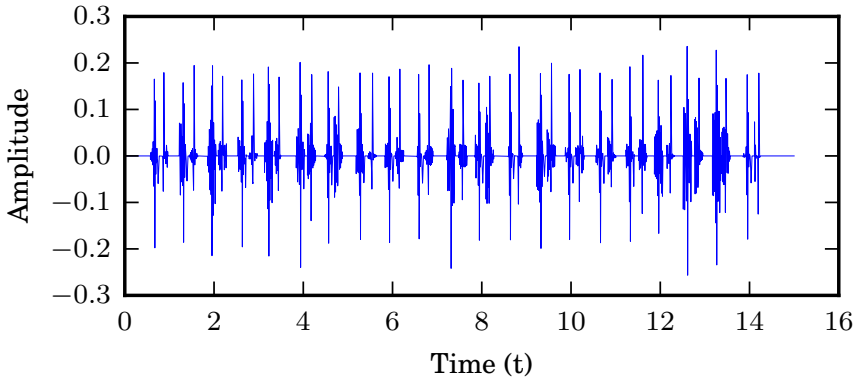

Figure 5.9: Graph showing a generated signal with correlated noise with a SNR of 5dB.

The results for the testing at three different heart rates are shown in Figures 5.10, 5.11, and 5.12.

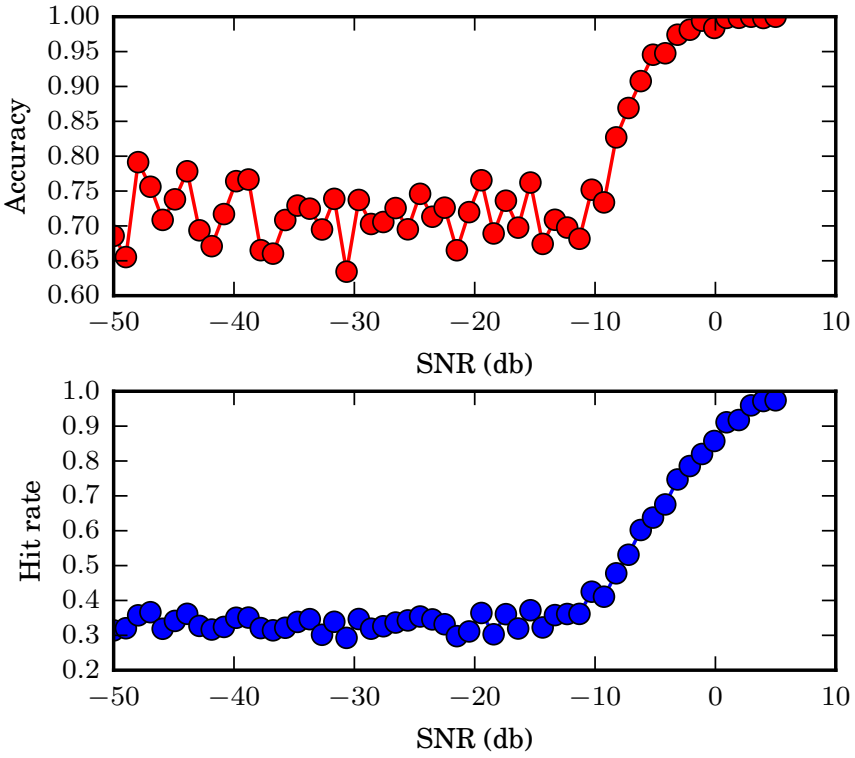

Figure 5.10: Graph showing the relationship between SNR and performance at 60 BPM.

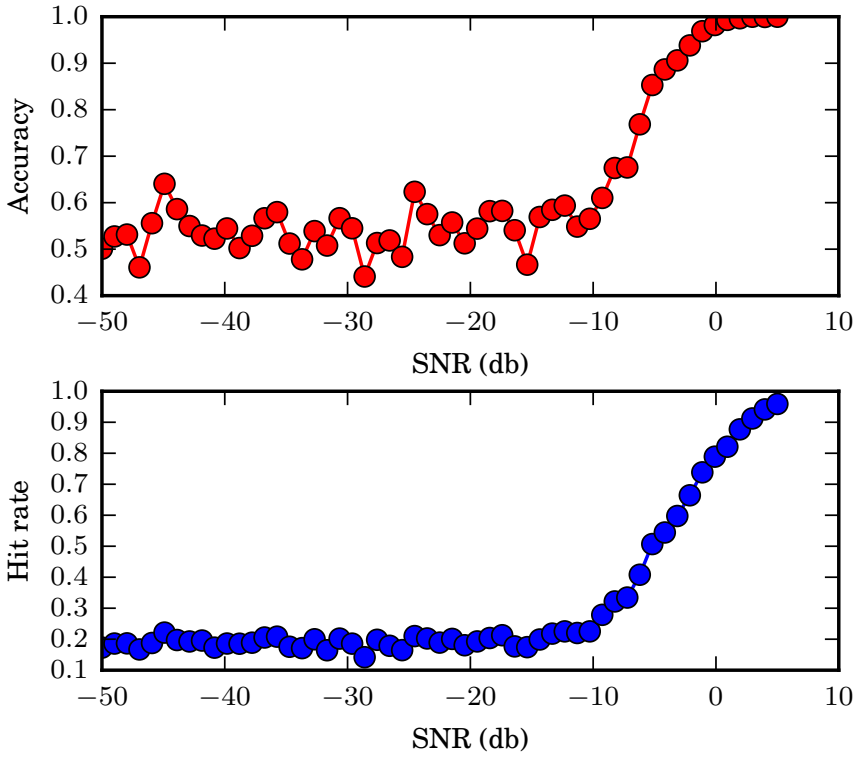

Figure 5.11: Graph showing the relationship between SNR and performance at 90 BPM.

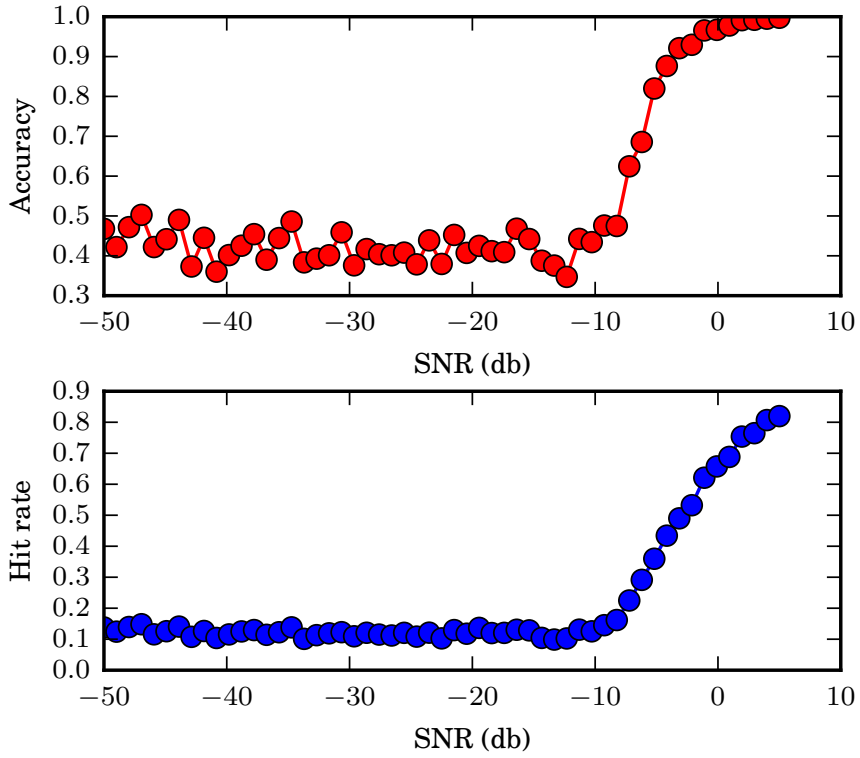

Figure 5.12: Graph showing the relationship between SNR and performance at 120 BPM.

5.2.4 Correlated and uncorrelated noise testing results

Four examples of signals with correlated and white noise are provided in Figures 5.13, 5.14 and 5.15

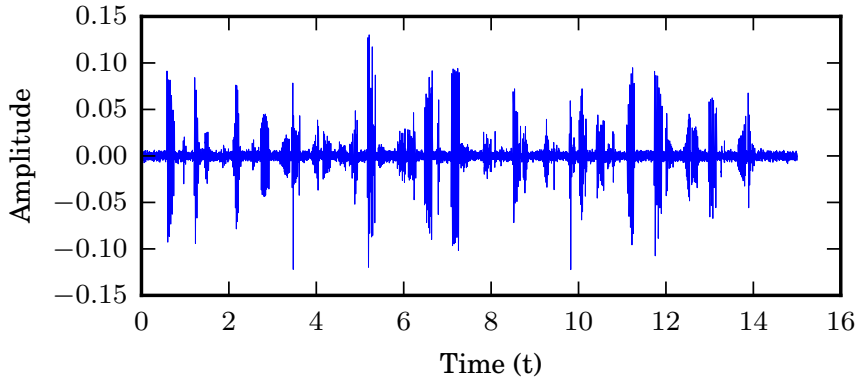

Figure 5.13: Graph showing a generated signal with correlated noise with a SNR of -50dB.

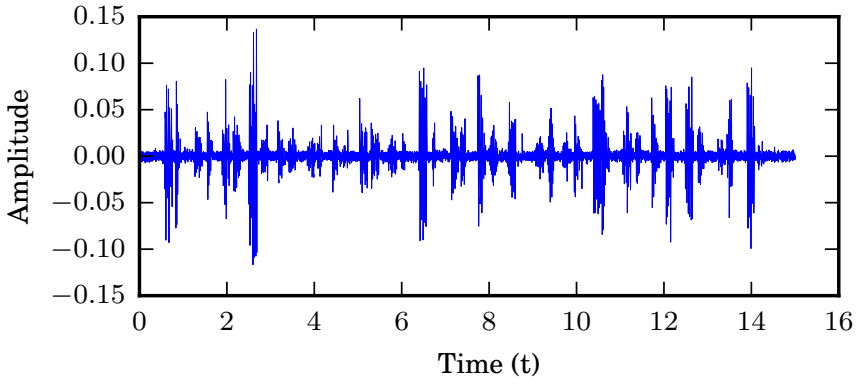

Figure 5.14: Graph showing a generated signal with correlated noise with a SNR of -15dB.

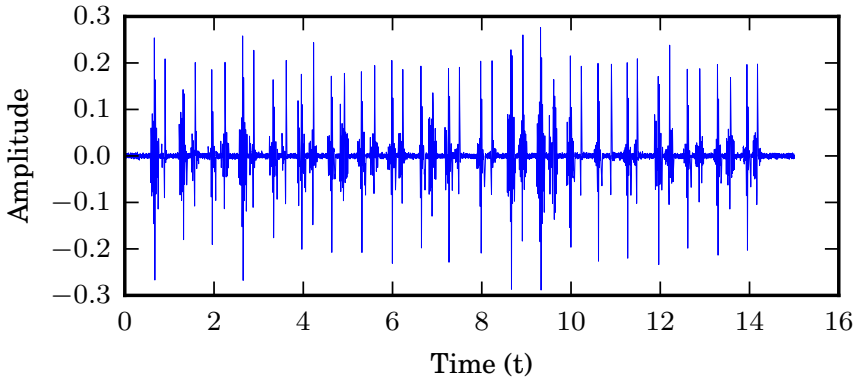

Figure 5.15: Graph showing a generated signal with correlated noise with a SNR of 5dB.

The results for the testing at three different heart rates are shown in Figures 5.4, 5.5, and 5.6.

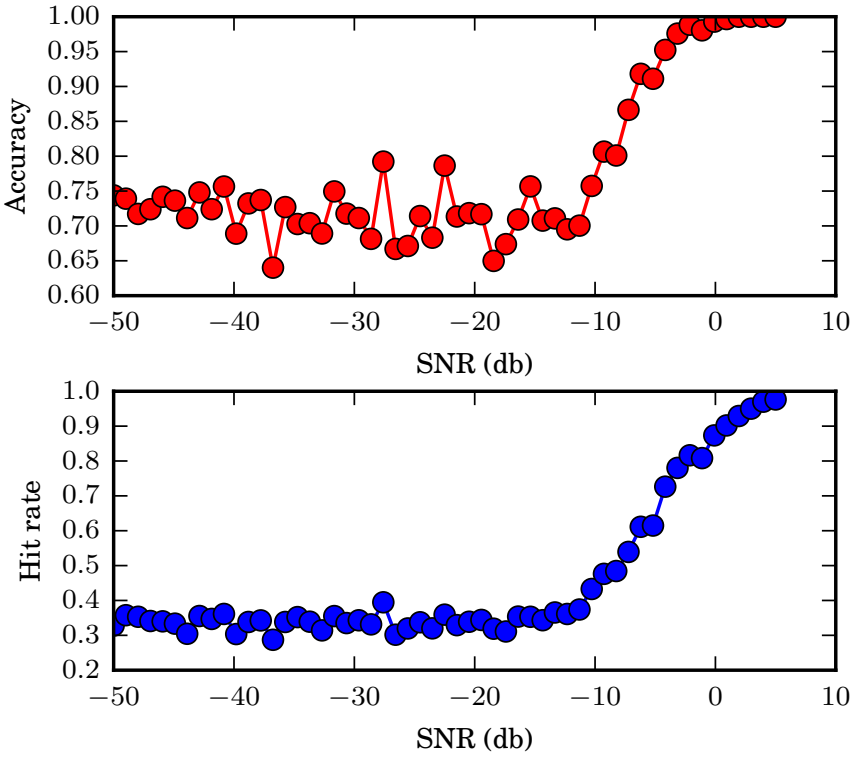

Figure 5.16: Graph showing the relationship between SNR and performance at 60 BPM.

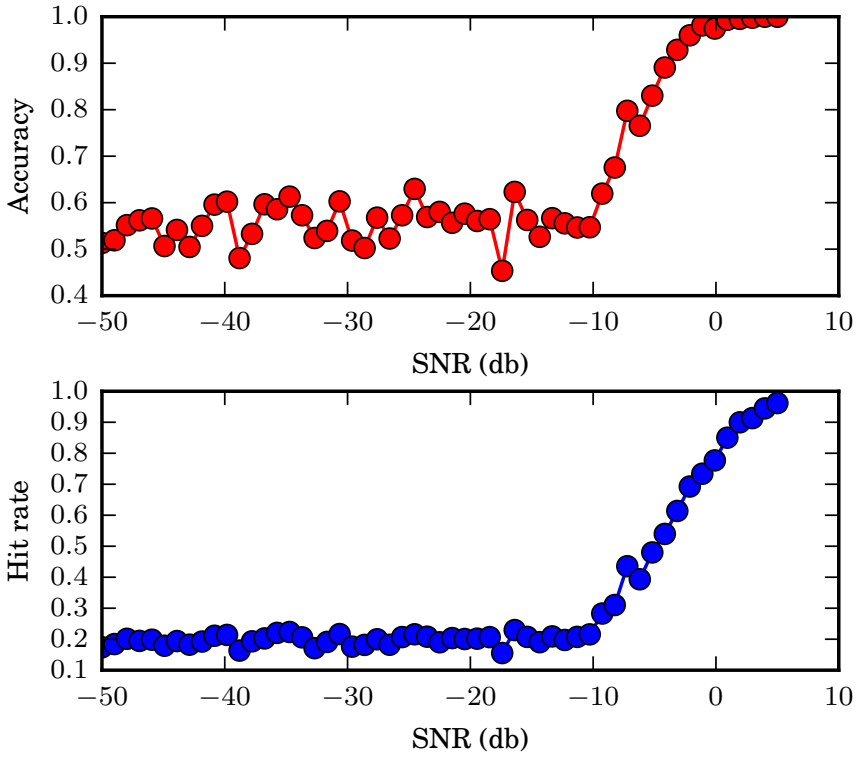

Figure 5.17: Graph showing the relationship between SNR and performance at 90 BPM.

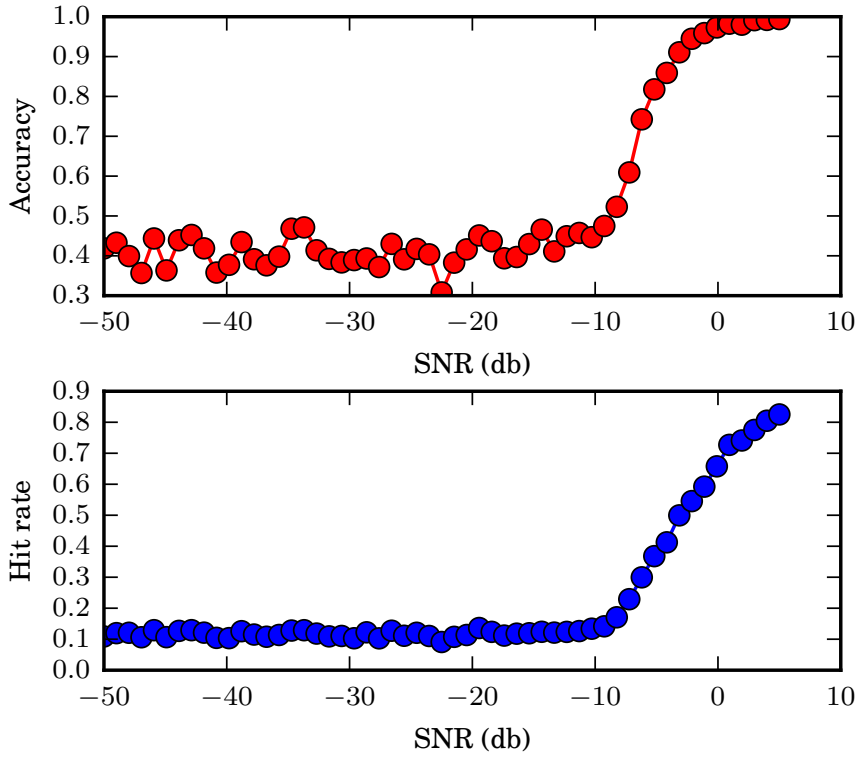

Figure 5.18: Graph showing the relationship between SNR and performance at 120 BPM.

5.3 Online testing

The functionality of the real time system was tested by monitoring the heart sounds of the author. The screen capture of the process in shows how the system is capable of segmentation in real time. The video can be viewed on the link in [34]. The video shows 60 seconds of real time analysis in which 39 pairs of S1 and S2 sounds are observed. Out of these 39 pairs, 24 were identified as S1-S2 pairs, with no incorrect sounds reported. This preliminary test shows the system is capable of identifying heart sounds in real time.

5.4 Summary

This section shows that the algorithm described in this thesis is capable of accurate heart sound segmentation of real signals, both in an online and offline capability. The algorithm is robust to different types of noise as the performance of the algorithm is more dependent on the amount of noise than the type. This makes it suitable for "real world" conditions where murmurs are present.

Chapter 6

Conclusion

From the examination of the literature, it was postulated that it was possible to identify a heart sound from other heart sounds based on its time-frequency content, which would satisfy Hypothesis 1.1 as time frequency content is short in length. The development and subsequent testing of the offline algorithm showed that the time frequency content of individual heart sounds, when transformed and correlated provides a feature vector that can separate the different types of heart sounds. An average accuracy of 84.2 % and a hit rate of 84.4 % indicate that the algorithm is a suitable candidate for further development. The algorithm in this thesis was tested against synthetic heart sounds that have a known signal to noise ratio. This allowed for the determination of where the algorithm will start to fail. The upper bound for where the algorithm will start to fail is -10dB of noise for the three heart rates tested in this thesis. The two offline testing results prove Hypothesis 1.1 and 1.2.

The real time segmentation of the heart sounds was tested on a single subject and was found to work flawlessly in detecting S1 and S2. This result and the short feature vector calculation time show prove Hypothesis 2.1.

Future work should be aimed at exploring the different types of time frequency analysis methods and testing them with the procedures in this thesis, along with other segmentation algorithms. The integration of existing methods of analysing segmented sounds is also a potential candidate for future work. Given the financial constraints of the expected operating environment, an efficient mobile implementation of the algorithm should also be investigated for use in telemedicine.

Bibliography

[1] Nannan, N., Dorrington, R., Laubscher, R., Zinyakatira, N., Prinsloo, M., Darikwa, T., Matzopoulos, R. and Bradshaw, D.: Under-5 mortality statistics in south africa. 2012.

[2] Martini, F.H., Bartholomew, E.F., Ober, W.C., Garrison, C.W., Welch, K. and Hutchings, R.: *Essentials of anatomy & physiology*. Pearson, 2003.

[3] Hoffman, J.I. and Kaplan, S.: The incidence of congenital heart disease. *Journal of the American College of Cardiology*, vol. 39, no. 12, pp. 1890–1900, 2002.

[4] Shub, C.: Echocardiography or auscultation? how to evaluate systolic murmurs. *Canadian family physician*, vol. 49, no. 2, pp. 163–167, 2003.

[5] Thompson, W., Hayek, C., Tuchinda, C., Telford, J. and Lombardo, J.: Automated cardiac auscultation for detection of pathologic heart murmurs. *Pediatric cardiology*, vol. 22, no. 5, pp. 373–379, 2001.

[6] Geggel, R.L., Horowitz, L.M., Brown, E.A., Parsons, M., Wang, P.S. and Fulton, D.R.: Parental anxiety associated with referral of a child to a pediatric cardiologist for evaluation of a still's murmur. *The Journal of pediatrics*, vol. 140, no. 6, pp. 747–752, 2002.

[7] Watrous, R.L., Thompson, W.R. and Ackerman, S.J.: The impact of computer-assisted auscultation on physician referrals of asymptomatic patients with heart murmurs. *Clinical cardiology*, vol. 31, no. 2, pp. 79–83, February 2008. ISSN 0160-9289.
Available at: `http://www.ncbi.nlm.nih.gov/pubmed/18257026`

[8] Zimetbaum, P. and Goldman, A.: Ambulatory arrhythmia monitoring choosing the right device. *Circulation*, vol. 122, no. 16, pp. 1629–1636, 2010.

[9] Coovadia, H., Jewkes, R., Barron, P., Sanders, D. and McIntyre, D.: The health and health system of south africa: historical roots of current public health challenges. *The Lancet*, vol. 374, no. 9692, pp. 817–834, 2009.

[10] Hoosen, E.G.M., Cilliers, A.M., Hugo-Hamman, C.T., Brown, S.C., Lawrenson, J.B., Zuhlke, L. and Hewitson, J.: Paediatric cardiac services in South Africa. *SAMJ: South African Medical Journal*, , no. 2, p. 106. ISSN 0256-9574.

[11] Hanna, I.R. and Silverman, M.E.: A history of cardiac auscultation and some of its contributors. *The American journal of cardiology*, vol. 90, no. 3, pp. 259–267, 2002.

[12] Felner, J.M.: The First Heart Sound. In: Walker, H., WD, H. and JW, H. (eds.), *Clinical Methods: The History, Physical, and Laboratory Examinations.*, 3rd edn, chap. 22. Butterworths, 1990.

[13] Alam, U., Asghar, O., Malik, R.A., Khan, S. and Hayat, S.: Cardiac auscultation: an essential clinical skill in decline. *The British Journal of Cardiology*, , no. 1, pp. –10.

[14] Dice, J.E. and Bhatia, J.: Patent ductus arteriosus: an overview. *The Journal of Pediatric Pharmacology and Therapeutics*, vol. 12, no. 3, pp. 138–146, 2007.

[15] Pretorius, E., Cronje, M.L. and Strydom, O.: Development of a pediatric cardiac computer aided auscultation decision support system. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, vol. 2010, pp. 6078–82, January 2010. ISSN 1557-170X.
Available at: `http://www.ncbi.nlm.nih.gov/pubmed/21097128`

[16] Felner, J.M.: The Second Heart Sound. In: Walker, H., Hall WD and Hurst JW (eds.), *Clinical Methods: The History, Physical, and Laboratory Examinations.*, 3rd edn, chap. 23. Butterworths, 1990.
Available at: `http://www.ncbi.nlm.nih.gov/books/NBK341/`

[17] Silverman, M.E.: The third heart sound. In: Walker, H., WD, H. and JW, H. (eds.), *Clinical Methods: The History, Physical, and Laboratory Examinations.*, 3rd edn, chap. 24. Butterworths, 1990.

[18] Williams, E.S.: The fourth heart sound. In: Walker, H., WD, H. and JW, H. (eds.), *Clinical Methods: The History, Physical, and Laboratory Examinations.*, 3rd edn, chap. 25. Butterworths, 1990.

[19] Geva, T., Martins, J.D. and Wald, R.M.: Atrial septal defects. *The Lancet*, vol. 383, no. 9932, pp. 1921–1932, 2014.

[20] Penny, D.J. and Vick, G.W.: Ventricular septal defect. *The Lancet*, vol. 377, no. 9771, pp. 1103–1112, 2011.

[21] Baumgartner, H.: Aortic stenosis: medical and surgical management. *Heart*, vol. 91, no. 11, pp. 1483–1488, 2005.

[22] Bekeredjian, R. and Grayburn, P.A.: Valvular heart disease aortic regurgitation. *Circulation*, vol. 112, no. 1, pp. 125–134, 2005.

[23] Choi, S. and Jiang, Z.: Comparison of envelope extraction algorithms for cardiac sound signal segmentation. *Expert Systems with Applications*, , no. 2, pp. 1056–1069, February. ISSN 09574174.

[24] Oskiper, T. and Watrous, R.: Detection of the first heart sound using a time-delay neural network. In: *Computers in Cardiology*, pp. 537–540. IEEE, 2002. ISSN 0276-6547.

[25] Kumar, D., Carvalho, P., Antunes, M., Henriques, J., Eugenio, L., Schmidt, R. and Habetha, J.: Detection of S1 and S2 heart sounds by high frequency signatures. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, vol. 1, pp. 1410–6, January 2006. ISSN 1557-170X.
Available at: `http://www.ncbi.nlm.nih.gov/pubmed/17946890`

[26] Ricke, A.D., Povinelli, R.J., Johnson, M.T. and Healthcare, G.E.: Automatic Segmentation of Heart Sound Signals Using Hidden Markov Models. pp. 953–956, 2005.

[27] Papadaniil, C.D. and Hadjileontiadis, L.J.: Efficient Heart Sound Segmentation and Extraction Using Ensemble Empirical Mode Decomposition and Kurtosis Features. *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 4, pp. 1138–1152, July 2014.

[28] Sejdić, E. and Jiang, J.: Pattern Recognition in Time-Frequency Domain: Selective Regional Correlation and Its Applications. In: Peng-Yeng, Y. (ed.), *Pattern Recognition Techniques, Technology and Applications*, chap. 27, pp. 13–626. I-Tech, Vienna, 2008. ISBN 978-953-7619-24-4.

[29] Arvin, F. and Doraismay, S.: Real-time segmentation of heart sound pattern with amplitude reconstruction. In: *2010 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, pp. 130–133. IEEE, November. ISBN 978-1-4244-7599-5.

[30] Northrop, R.B.: *Signals and systems analysis in biomedical engineering*. CRC press, 2010.

[31] Gupta, C.N., Palaniappan, R., Swaminathan, S. and Krishnan, S.M.: Neural network classification of homomorphic segmented heart sounds. *Applied Soft Computing*, , no. 1, pp. 286–297, January. ISSN 15684946.

[32] Stein, E.M. and Shakarchi, R.: *Fourier analysis: an introduction*, vol. 1. Princeton University Press, 2011.

[33] Hartigan, J.A. and Wong, M.A.: Algorithm AS 136: A K-Means Clustering Algorithm. *Applied Statistics*, vol. 28, no. 1, p. 100, January 1979. ISSN 00359254.

[34] Fourie, D.: Real time segmentation of heart sounds. 2015.
Available at: `https://www.youtube.com/watch?v=W9M6rh134p4`